# Columbia University

# Contributions to Education

## Teachers College Series

# HOW TO CONSTRUCT THE TRUE-FALSE EXAMINATION

BY

CHARLES CONRAD WEIDEMANN, PH.D

ASSOCIATE PROFESSOR OF EDUCATION, UNIVERSITY OF NEBRASKA
LINCOLN, NEBRASKA

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
CONTRIBUTIONS TO EDUCATION, NO. 225

BUREAU OF PUBLICATIONS
Teachers' College, Columbia University
NEW YORK CITY
1926

# ACKNOWLEDGMENTS

A grant from the Commonwealth Fund to the Bureau of Collegiate Educational Research of Columbia College made this study possible.

To Dean W. W. Kemp and his associates of the University of California is given credit for a satisfying introduction to service in education.

Indebtedness is due Professors N. L. Engelhardt, P. R. Mort, C. Alexander, F. H. Swift, and F. W. Hart for coöperation and assistance before this study was undertaken and during its progress.

The author wishes to thank a group of student associates for their assistance. Mr. Paul D. Collier, Mr. Willard S. Ford, and Mr. Roger M. Thompson timed and recorded the results of Chapter V. Mr. Lester W. Bartlett, Mr. Oscar W. Irvin, Mr. Hubert H. Smith, and Mrs. E. P. W. Weidemann judged the experimental procedures for Chapters I, II, and IV. Dr. John R. Clark also assisted in this preliminary step.

Appreciation is due to Dr. John P. Herring, Dr. William A. McCall, Dr. Percival M. Symonds, and Dr. Ben D. Wood for their coöperation and expert judgment of the experiments described in Chapters I, II, and IV.

He wishes to express gratitude to the members of his dissertation committee. He is grateful to Professors W. F. Russell and E. S. Evenden for constructive suggestions; to Professor J. Ralph McGaughy for assistance in procedures and form; to Professor Ben D. Wood for many keen and discriminating criticisms; and to Professor George Drayton Strayer, chairman, for insight, inspiration, and guidance during this study.

The author desires to express his appreciation of the encouragement, help, and faithfulness, often to the point of obvious sacrifice, which his wife, E. Pearl Walther Weidemann, has given.

To his mother and father, still at the hearth stone, go the deepest gratitude and thankfulness for the urge to "carry on."

C. C. W.

# CONTENTS

# HOW TO CONSTRUCT THE TRUE-FALSE EXAMINATION

## INTRODUCTION

### STATEMENT OF THE PROBLEM AND ITS PURPOSE

The purpose of this study is to demonstrate how the true-false statement may be improved. The two major divisions with which improvement is concerned are (1) administrative conveniences and (2) content. Administrative conveniences include for the greater part those factors which (1) aid the student to read and to record his response, and (2) contribute to speed and accuracy of scoring. Content deals with the nature of the subject matter which is related to or contained in any given item or statement of an examination.

The study is intended to be of assistance to the classroom instructor from about the sixth grade in the elementary schools through the university. Administrators and test constructors should also find the suggestions useful.

#### IMPORTANCE OF EXAMINATIONS

With respect to examinations James E. Russell [1] says:

> Such tests as written recitations, quizzes, term and final examinations, and the like are of the greatest value to the teacher who is really concerned in education of his pupils. The examinations are indispensable; they need no argument to justify the position they hold in our scheme of instruction.

Then again Dean Russell speaks of the evolution of a unified and really American system of Education made up of nicely adjusted constituent parts.[2] He says:

> When that time comes I venture to predict we shall hear nothing of examination for admission to any grade or to any school, but *much will be said of examination for instruction and promotion.*

[1] Russell, James E. *The Trend in American Education*, p. 48.
[2] *Ibid.*, p. 57.

He sums up a problem for solution, as follows: [3]

Examinations must have a place in every scheme of instruction. Instruction can proceed only when the extent and quality of the learner's knowledge is *definitely* understood. Every recitation, every review is such an examination; further examinations of a formal sort are often desirable for the sake both of the teacher and of the pupil. But such examinations are given by the teachers within the school or school system and primarily for the purpose of instruction.

Again, George D. Strayer and N. L. Engelhardt say: [4]

Teachers have always sought to measure the achievement of their pupils. . . . In our practice we have tended to confine our measurements largely to examinations which test knowledge, memory, or habit. These examinations . . . have been the basis upon which they have moved from grade to grade in a school or transferred from one type of school to another.

Teachers have long recognized the difficulty of comparing results upon the basis of examinations. Even though the same general division or part of a field has been the subject of an examination the tests which have been set varied so greatly that it has been practically impossible to compare children from year to year or from different school systems.

And again, they say: [5]

In any field the development of units of measurement is dependent upon careful investigation and upon a realization of the imperfections of the units already used. . . . The more imperfect the unit of measure which we now apply, the greater the necessity for insisting upon measurement.

## DEFINITIONS

A true-false statement is a declarative sentence which is either *true* or *false* (not *true*) according to a criterion of acceptability. The criterion of acceptability for this study is any group of three or more instructors in a given subject matter course who approximately agree that a given statement is definitely either *true* or *false*. A more rigid criterion based upon the behavior of a validating group of students in the subject matter is not of primary concern in this study. It is maintained that the problem of improvement of the content of true-false statements upon the behavioristic basis need not concern the classroom instructor. This problem is more suited for treatment in a Bureau of Educational Research.

The instructor may be given a series of "rules" whereby he

[3] Russell, James E. *The Trend in American Education*, pp. 58-59.
[4] Strayer, George D., and Engelhardt, N. L. *The Classroom Teacher*, p. 165.
[5] *Ibid.*, p. 198.

may, if he desires, improve his examination statements in many concrete ways. Among other things this study attempts to assemble a few of the more helpful "rules." (For concrete illustrations of the term "rules" see Chapter VII.)

### CRITERION OF IMPROVEMENT

Improvement assumes that true-false statements exist which are susceptible to change in a direction of from "not good" to "good." A behavioristic criterion of "not good" and "good" for true-false statements is quoted in Chapter X from a recent publication.[6]

### SOURCES OF DATA

A total of 17,047 true-false statements was received from 175 examinations in 48 general subject matter fields from 48 colleges and universities distributed over the United States. About 6000 of the 17,047 statements were received from the Bureau of Collegiate Research of Columbia College, Columbia University. These data are of May and June, 1925.

TABLE 1

DISTRIBUTION OF THE TRUE-FALSE EXAMINATIONS USED IN THIS STUDY BY SUBJECT MATTER GROUPS

| Subject Matter* Group Number | | Number of True-False Examinations |
|---|---|---|
| I. | Languages | 19 |
| II. | Social Sciences | 78 |
| III. | Natural Sciences | 15 |
| IV. | Physical Sciences | 11 |
| V. | Engineering Sciences | 6 |
| VI. | Applied Sciences | 33 |
| VII. | Psychology and Philosophy | 10 |
| VIII. | Fine Arts | 3 |
| | Total | 175 |

*I. Languages: English, French, Spanish, German.
II. Social Sciences: History, Economics, Government, Political Science, Contemporary Civilization, Sociology, Law, Education, Educational Psychology.
III. Natural Sciences: Botany, Zoölogy, Biology, Geography, Geology.
IV. Physical Sciences: Chemistry, Physics, Astronomy, Mathematics.
V. Engineering Sciences: Mechanical, Electrical, Mining, Civil.
VI. Applied Sciences: Medical Courses, Agricultural Courses, Commercial Courses, Physical Education, Home Economics.
VII. Psychology and Philosophy.
VIII. Fine Arts: Industrial and Applied Art, Fine Arts.

[6] Wood, Ben D. "Measurement of Law School Work, II," *Columbia Law Review*, Vol. XXV, No. 3, March, 1925.

Table 2 shows the distibution by geographical sections [7] of the 48 colleges and universities which provided true-false examinations.

TABLE 2

| Geographical Section | Number of Colleges and Universities in Each Section |
|---|---|
| Eastern | 18 |
| Southern | 9 |
| Great Lakes | 10 |
| Great Plains | 4 |
| Western Section | 6 |
| No Name | 1 |
| Total | 48 |

### AVERAGE LENGTH OF AN EXAMINATION

The average length of an examination was 97 statements. The range was from 1 to 244 statements to an examination. The examination papers were assembled from the instructors in 48 colleges and universities in the United States.

The courses in Educational Administration and Philosophy of Education given in Teachers College are represented by one examination each. These data are of December, 1925, and January, 1926, respectively.

The Achievement Intelligence tests published by (1) The Bureau of Publications of Teachers College, Columbia University, (2) The Public School Publishing Company, Chicago, and (3) The World Book Company, Yonkers-on-Hudson, contained a small additional amount of data. These data include publications up to January, 1925.

### METHODS OF THE STUDY

Analysis was used to reveal the limits and details of the problem.

Experimental procedures were used to ascertain the quantitative and qualitative relationships within the problem.

Statistical procedures were used to reveal the significance of the relationships within the problem.

[7] McGaughy, J. Ralph. *The Fiscal Administration of City School System.*

GENERAL PROCEDURE

The problem has developed through a sequence of five steps.

1. Analysis of Administrative Conveniences and Content.
2. Experiments of Administrative Conveniences and Content.
3. Statistical treatment of the results for significance.
4. Interpretation of the results.
5. Application of the results to subject matter through concrete examples.

# PART ONE

# ADMINISTRATIVE CONVENIENCES

## CHAPTER I

### DIRECTIONS FOR THE EXAMINATION

#### PRESENT PRACTICE

It is not possible to include all the 227 sets of Directions used in this study. In some cases the same set was used with different examinations. It should serve the purpose of this study to present typical samples of what has actually been selected as representative of a range of their content. The samples were selected by the author. No claim is made that worse or better samples might not have been included.

The samples below should be considered independent of the variations of how the statement is to be marked or where the mark is to be placed. *How* and *Where* are essential abstractions in the content of Directions. The concrete ways in which the *How* or *Where* is executed are not the fundamental bases for analysis of content of Directions.

The samples are exact reproductions, including typographical errors.[1]

Sample 2.

If what the statement below says is TRUE, put a plus sign (+) in column A; if what it says is FALSE, put a minus sign (—) in column A. Mark every statement; a statement not marked will be scored wrong.

ASK NO QUESTIONS. DO NOT QUALIFY ANY OF THE STATEMENTS.

Sample 4.

(This sample represents the case where no DIRECTIONS have been given.)

[1] Complete data are given on pages 9-10 of manuscript copy of this dissertation, Teachers College Library, Columbia University.

Sample 6.

Answer "YES" or "NO."

Sample 9.

False-True.

Sample 11.

Mark *X in front* of statement if all or any part of it is *wrong* or *incomplete*. If right or complete do not mark.

Sample 13.

Approximately half of the statements incline to be true and the remainder false. Mark each T (true) or F (false), or U if undecided. Place letted *before* statement.

Sample 14.

Chech with an X those statements that are incorrect.

Sample 19.

To denote your opinion, write True or False, as the case may be, after each statement.

Sample 22.

Write "yes" in front of each statement that is generally true and O in front of those generally false.

Sample 24.

True-false test. Encircle T or F. Restate false or ambiguous statements in your blue-book.

Sample 25.

Mark each statement true or false and make a brief statement in your blue-book justifying yout judgment. Refer to the statements by number, and save rewriting them.

The extremely wide range of variability of statements is at once recognizable.

## STATEMENT OF PURPOSES

The purposes of this chapter are:

1. To construct sets of Directions by expert judgment [2] from the materials of Present Practice.

2. To state the factors of agreement and disagreement of judgments in the sets of Directions.

3. To compare the sets of Directions for frequency of use of words and sentence structures.

[2] Five judges were selected to judge the materials in Chapter I, "Directions for the Examination;" Chapter II, "Indication of the Response;" and Chapter IV, "Assembling the Statements." The four experts were selected with the help of Dr. John R. Clark, of Lincoln School, Teachers College, Columbia University.

The writer served as one of the five judges and recorded his judgment prior to that of the four experts. The four expert judges were: Dr. John P. Herring, Research Associate, Institute of Child Welfare Research, Speyer School, Teachers College, Columbia University; Dr. William A. McCall, Associate Professor of Education, Teachers College; Dr. Percival M. Symonds, Assistant Professor of Education, Teachers College; Dr. Ben D. Wood, Director of the Bureau of Collegiate and Educational Research, Columbia College, Columbia University.

4. To express a relationship between the length of a set of Directions and the frequency of their use in a true-false examination.

### CONSTRUCTION OF THE SETS OF DIRECTIONS

*Procedure.*—1. An analysis of the 227 sets of Directions for unit sentences revealed a possibility of classification under seven main divisions:

1 unit sentence for *kind* of statement
3 unit sentences for *how* and *where* to indicate the response
1 unit sentence for *scoring information*
1 unit sentence for *general procedure*
1 unit sentence for *general behavior*
7 unit sentences in all

2. Ten of what the writer considered samples of the worst to better and best unit simple declarative or imperative sentences contained in Directions as a basis to condition and guide a student's reactions correctly, were assembled under each of the seven main divisions. This made a total selection of seventy unit sentences from the 227 sets of Directions. These seventy unit sentences have been classified into seven packets of ten 5 x 8 inch cards in each packet.[3] (See Appendix A, Materials of the Experiment.)

3. A carefully planned memorandum to the expert judges was prepared. The memorandum specified the procedure for judgment of the materials in the Packets. In addition, each judge was asked to improve upon the unit sentences of the Packets wherever he considered it possible. Thus each judge prepared a set of Directions independent of the others.

As a final step, each expert judge was given a typewritten copy of each of the five sets of Directions, and was asked to formulate a final set of Directions in careful detail.[4]

[3] The above sampling of the Directions was judged to be approximately complete. The words true-false were used. The words and symbols for plus and zero were used. The elimination of variation of mere names, like Right-Wrong, Yes-No, Correct-Incorrect, etc., and mere varieties of ways of marking a statement simplifies the problem to an extent that makes possible a rather extensive sampling of sentence structure. In a few cases it became necessary to invent unit sentences. The sets of Directions at hand did not contain enough varieties of statement to give ten samples under each heading. The *kind of statement, how to mark, where to place the mark*, unit sentences are all actually taken from the total sample of 227 sets of Directions.

[4] Full details of procedure are explained on pages 11-16, 218-22, 229 of manuscript copy of this dissertation, Teachers College Library.

*Results.*—The following is the final set of directions formulated by each of the five judges.

SETS OF DIRECTIONS[5]

Set No. 1

Each of the statements below is either true or false. If a statement is TRUE, put a plus sign, +, on the dotted line to the left of the *true* statement. If a statement is partly or entirely FALSE, put a zero, 0, on the dotted line to the left of the *false* statement. Do not mark a statement for which you have no decision.

DO NOT GUESS. Guessing reduces your score.

Do all the easy ones first, then return and do the harder ones.

Do not talk or ask questions.

Set No. 2

Some of the following statements are *true* and some are *false*. Indicate all statements that are *true* by a plus sign. Indicate all statements that are *false* by a zero. Put the plus sign or the zero on the dotted line.

In case of doubt guessing is permitted, but a wrong guess reduces your score more than an omission.

Ask no questions during the examination.

Set No. 3

Put a + on the dotted line before the TRUE statement and a 0 on the dotted line before the FALSE statement.

Mark the statements in order.

Omit no statements.

If you do not know, guess.

Ask no questions.

Set No. 4

Some of the following statements are true and some are false. If a statement seems *true,* put a + on the dotted line before the true statement. If a statement seems false, put a 0 on on the dotted line before the false statement.

Omit no statements. If you do not know, guess.

Do not talk or ask questions during the examination.

Set No. 5

About one-half of the following statements are *true* and about one-half are *false*. Mark each *true* statement with a plus sign (+) on the

[5] For packet card numbers of preliminary sets of Directions see Chart 1, page 15 of manuscript copy of this dissertation, Teachers College Library.

dotted line at the left of the statement. Mark each statement that is partly or wholly *false* with a zero (0) on the dotted line at the left of the statement. Do not mark statements which you either do not know or which you judge to be ambiguous.

Answer all the easy ones first, then return to the harder ones, if you have time.

Ask no questions; if in doubt, read the directions again and use your own judgment.

DO NOT GUESS. A wrong answer decreases your score more than an omission. Your score is the number that you answer correctly minus the number that you answer incorrectly. Omissions are not subtracted from the number that you answer correctly.

### RELATIONSHIP OF THE LENGTH OF DIRECTIONS TO THE FREQUENCY OF THEIR USE IN A TRUE-FALSE EXAMINATION

It is the agreed opinion of the four experts who assisted in this study that the Directions for a series of true-false statements should be varied in length to fit the experience, age, and intelligence of the students.

If a true-false examination is to be given to a group of students who never saw a series of true-false statements before, it seems reasonable to use a more detailed set of Directions than a very short set of Directions.

One might use Directions Set No. 1 the first time a group of students ever took a true-false examination. One year later Directions Set No. 3 probably would be satisfactory. In another year the same students might easily proceed without any difficulty whatever with a set of Directions which contained a bare sentence or two on *how* to mark, and *where* to place the mark of a true-false statement. Even more abbreviated forms are conceivable as satisfactory with certain well-established, long-standing habits of taking true-false examinations.

When the limit of brevity is approached the words *True* (+) and *False* (0), with a dotted line before each statement might satisfactorily serve the purpose of an instructor. Such brevity of Directions would assume that the students to whom the examination was to be given had taken many true-false examinations before.

The unvalidated statement of the principle involved is: Use of true-false examinations probably permits gradual shortening of the Directions preceding such examinations.

EXAMPLES IMMEDIATELY FOLLOWING A SET OF DIRECTIONS . . .

When a true-false examination is to be given for the first time to a group of students, every precaution should be observed to insure complete and full comprehension of Directions upon the part of each student. This unusual precaution becomes less and less necessary as the students become accustomed to the procedure of taking true-false examinations.

The *example* is a device which serves to illustrate concretely every kind of reaction which the student may encounter during the examination.

To illustrate: Directions Set No. 1 states *how* true statements and *how* false statements are to be marked, also the place *where* the mark is to be placed specified, and so on.

Immediately below the Directions the following examples might be placed:

EXAMPLES:

TRUE———+ 1. Cæsar was a Roman.
FALSE———0 2. Cæsar was a Greek.
OMISSION— 3. Cæsar loved to fish.

Such precaution seems unnecessary when students have taken a few true-false examinations.

## COMPARISON OF THE SETS OF DIRECTIONS FOR FREQUENCY OF USE OF WORDS AND SENTENCE STRUCTURES

*Procedure.*—1. Frequency distributions were constructed of the words in the five sets of Directions according to the Thorndike Word Book of the 10,000 most frequently used words in the English language.

2. Frequency distributions were constructed of the sentence structures of the five sets of Directions.

3. The number and kind of dependent clauses were also listed.

*Results.*—Table 3 indicates that more than fifty per cent of all of the words in the five sets of Directions are among the first hundred most frequently used words in the English language.[6] More than ninety per cent of all of the words are among the first

[6] Thorndike, E. L. *The Teacher's Word Book*, p. 134. Teachers College, Columbia University, 1921.

five thousand most frequently used words in the English language. The only words in the five sets of Directions that occur between the 5,000 and 10,000 points of frequency of occurrence of words were: *zero, plus, minus, ambiguous,* and *subtracted.*

TABLE 3

FREQUENCY DISTRIBUTIONS OF WORDS IN THE FIVE SETS OF DIRECTIONS

| Frequency Placement | Judges | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | No. 1 | | No. 2 | | No. 3 | | No. 4 | | No. 5 | |
| | Frequency | Per Cent | Frequency | Per Cent | Frequency | Per Cent | Frequency | Per Cent | Frequency | Per Cent |
| 1 *a* 1.. | 48 | 54.0 | 34 | 52.3 | 20 | 51.3 | 31 | 50.8 | 80 | 55.9 |
| 1 *a* 2.. | 6 | 6.7 | 1 | 1.5 | 1 | 2.6 | 2 | 3.3 | 9 | 6.3 |
| 1 *a* 3.. | 6 | 6.7 | 4 | 6.1 | 4 | 10.3 | 8 | 13.1 | 9 | 6.3 |
| 1 *a* 4.. | 3 | 3.4 | 1 | 1.5 | 2 | 5.1 | 2 | 3.3 | 9 | 6.3 |
| 1 *a* 5.. | 0 | .... | 0 | .... | 0 | ..... | 0 | .... | 0 | ..... |
| 1 *b*.... | 9 | 10.0 | 8 | 12.3 | 2 | 5.1 | 3 | 4.9 | 10 | 7.0 |
| 2 *a*.... | 1 | 1.1 | 1 | 1.5 | 0 | ..... | 0 | .... | 4 | 2.8 |
| 2 *b*.... | 6 | 6.7 | 6 | 9.2 | 3 | 7.7 | 5 | 8.2 | 7 | 4.9 |
| 3 *a*.... | 1 | 1.1 | 1 | 1.5 | 0 | ..... | 0 | .... | 0 | ..... |
| 3 *b*.... | 0 | .... | 1 | 1.5 | 0 | ..... | 1 | 1.6 | 0 | ..... |
| 4 *a*.... | 7 | 7.0 | 3 | 4.6 | 4 | 10.3 | 6 | 9.8 | 7 | 4.9 |
| 4 *b*.... | 0 | .... | 1 | 1.5 | 1 | 2.5 | 1 | 1.6 | 3 | 2.1 |
| 5 *a*.... | 1 | 1.1 | 2 | 3.1 | 1 | 2.5 | 1 | 1.6 | 1 | 0.7 |
| 5 *b*.... | 0 | .... | 0 | .... | 0 | ..... | 0 | .... | 0 | ..... |
| 6 *a*.... | 0 | .... | 0 | .... | 0 | ..... | 0 | .... | 0 | ..... |
| 5545 to 6047 | 0 | 1.1 | 0 | .... | 0 | ..... | 0 | .... | 1 | 0.7 |
| 6619 to 7262 | 1 | .... | 2 | 3.1 | 1 | 2.6 | 1 | 1.6 | 1 | 0.7 |
| 8146 to 9190 | 0 | .... | 0 | .... | 0 | ..... | 0 | .... | 1 | 0.7 |
| 9191 to 10000 | 0 | .... | 0 | .... | 0 | ..... | 0 | .... | 1 | 0.7 |
| Total. | 89 | 99.9 | 65 | 99.7 | 39 | 100.0 | 61 | 99.8 | 143 | 100.0 |

Legend
1 *a* 1 = word in the first hundred most frequently used words.
1 *a* 2 = word in the second hundred.
1 *b* = word in the second half of the first thousand.
2 *a* = word in the first half of the second thousand.
2 *b* = word in the second half of the second thousand.

FACTORS OF AGREEMENT AND DISAGREEMENT OF JUDGMENTS IN FIVE SETS OF DIRECTIONS

*Procedure.*—1. A comparison was made of the sets of Directions for factors of agreement and disagreement.

*Results.*—1. The five sets of Directions agree in the following respects:

*a.* A sentence expresses the kind of examination statement to be answered.

*b.* A sentence expresses *how to mark* and *where to place* the indication of the response to a *true* statement.

*c.* A sentence expresses *how to mark* and *where to place* the indication of the response to a *false* statement.

*d.* A sentence expresses the thought that questions should not be asked by the students during the examination.

2. The five sets of Directions disagree in the following respects:

*a.* Whether or not statements should be *omitted* by the student.

*b.* Whether or not students should *guess* those statements of which they have no information.

*c.* The order in which statements should be answered.

Whether or not students should *omit* or *guess* statements which they do not know resolves itself into the issue of GUESS or DO NOT GUESS. This is another problem for research.[7]

Table 4 indicates that simple sentences are seventy per cent of all of the sentences in the five sets of Directions.

TABLE 4

DISTRIBUTION OF THE SENTENCE STRUCTURES OF THE FIVE SETS OF DIRECTIONS

| Type of Sentence Structure | Directions Set No. | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | Total |
| Simple Sentence | 5 | 4 | 4 | 3 | 9 | 25 |
| Compound Sentence | 0 | 1 | 0 | 0 | 0 | 1 |
| Complex Sentence | 3 | 0 | 1 | 3 | 2 | 9 |
| Total | 8 | 5 | 5 | 6 | 11 | 35 |

[7] Wood, Ben D. "Studies of Achievement Tests." *Journal of Educational Psychology*, pp. 1-22, January, 1926.

The complex sentences are distributed between *condition* and *time* dependent clauses as follows:

| Connective | No. of Clauses |
|---|---|
| *if* | 7 |
| *then* | 2 |
| Total | 9 |

It is evident, according to the judgment of the four experts and the writer, that a set of Directions should consist of simple, clear, concise, and accurate sentences.

## CONCLUSIONS

The results of the judgments of the five expert judges indicated that:

1. The sets of Directions differ in length and word content.
2. The judges agree that a set of Directions should contain a sentence which expresses:
   - *a.* The *kind* of examination statement to be answered.
   - *b.* *How to mark* and *where to place* the response to a *true* statement.
   - *c.* *How to mark* and *where to place* the response to a *false* statement.
   - *d.* The thought that students should not ask questions during the examination.
3. The judges disagree with respect to:
   - *a.* Whether or not statements should be *omitted* by the student.
   - *b.* Whether or not students should *guess* the answer to those statements of which they have no information.
   - *c.* The order in which statements should be answered.
4. True-False Directions should consist of commonly used words.
5. True-False Directions should consist of simple sentences.
6. The length of directions probably should vary inversely with frequency of their use in a true-false examination.

# CHAPTER II

## THE INDICATION OF THE RESPONSE

### PRESENT PRACTICE

The methods of indication of a response to a true-false statement may be classified as (1) Freehand Written response, and (2) Mechanically Printed response.

In the case of Freehand Written responses the student is directed to write a word or a symbol standing for a word.

In the case of Mechanically Written responses the student is directed to *enscribe* (encircle or *enoval*), draw a line under or draw a line through a printed word or printed symbol standing for a word.

Various words and symbols have been used to indicate a response to a true-false statement. The samples below are representative of Present Practice.[1]

Freehand response—written symbols and written words.

R W, + −, + 0, R NR, T F, Yes No, Right Wrong, True False.

Mechanical response for either printed words or printed symbols are similar to the above samples.

### STATEMENT OF PURPOSE

The purpose of this chapter is to select certain words or symbols to be used by a student to indicate his response to each statement in a series of true-false statements.

*Procedure.*—1. The materials for the experiment were assembled into four Packets of twelve 5 x 8 inch cards. Each card contained one pair of either word opposites or symbolized word opposites similar to the above samples.[2]

2. A carefully planned memorandum was prepared and given to each of the expert judges. This memorandum specified the

[1] Complete data are given on pages 230-31 of manuscript copy of this dissertation, Teachers College Library.

[2] Complete list of materials is given on pages 240-44 of manuscript copy.

[3] For complete details of procedure see pages 27-29, 231-40 of manuscript copy of this dissertation, Teachers College Library.

procedure to be followed in the judgment of the materials in the Packets.[3]

*Results.*—Table 5 below contains a detailed account of the successive words and symbols which were selected by the expert judges.

TABLE 5

DISTRIBUTION OF THE *WRITTEN* AND *PRINTED* WORDS OR SYMBOLIZED WORDS FOR RESPONSES TO A TRUE-FALSE EXAMINATION

| Basis for Judgment | No. of the Operation | | Number of the Judge | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 |
| Freehand Written Response | 1 | Card No. | 9191 | 1749 | 1749 | 1749 | 9365 |
| | | Word | Plus | True | True | True | Yes |
| | | | Zero | False | False | False | No |
| | 2 | Card No. | 8459 | 8459 | 3165 | 8459 | 2579 |
| | | Symbol | +0 | +0 | t-f | +0 | +− |
| Mechanically Printed Response | 3 | Card No. | 0421 | 0421 | 0421 | 0421 | 0691 |
| | | Word | True | True | True | True | Yes |
| | | | False | False | False | False | No |
| | 4 | Card No. | 0876 | 0876 | 0876 | 0876 | 0876 |
| | | Symbol | t-f | t-f | t-f | t-f | t-f |
| Total Time Required to Write Response | 5 | Card No. | 8459 | 8459 | 3165 | 8459 | 2579 |
| | | Symbol | +0 | +0 | tf | +0 | +− |
| | 6 | Card No. | 0876 | 0876 | 0876 | 0421 | 0876 |
| | | | | | | 0876 | 0691 |
| | | Symbol | t-f | t-f | t-f | t-f | t-f |
| | | | | | | True | Yes No |
| | | | | | | False | |
| Symbols Easiest to Learn | 7 | Card No. | 8459 (write) | 0876 (circle) | 3165 (write) | 8459 (write) | 0876 (circle) 0691 (circle) 2579 (write) |
| | | Symbol | +0 | t-f | t-f | +0 | t-f |
| | | | | | | | Yes No |
| | | | | | | | +− |

Paterson [4] says: "The encircling of the letters 'T' or 'F' or the words 'True' or 'False' before the statements is perhaps the better method." (See Chapters V and VI for further evidence favoring the writing of the symbols + and 0.)

## CONCLUSIONS

The results of the judgments of the five expert judges indicate the following conclusions.

[4] Paterson, D. G. *Preparation and Use of New Type Examination*, p. 25.

1. The judges prefer symbolized word-opposites to word-opposites. (See Table 5, Operations Nos. 5 and 6.) The basis for this selection is time.

2. Operation No. 5 shows
   *a.* Four of the five judges favor the use of + as a response to a *true* statement.
   *b.* Three of the five judges favor the use of 0 as a response to a *false* statement.

3. Operation No. 7 is of doubtful value.

4. There is some reason to believe that + for *true* and 0 for *false* is to be preferred to indicate a response to a true-false statement. (See Chapters V and VI for further evidence.)

# CHAPTER III

## THE PROVISION FOR THE RESPONSE

### PRESENT PRACTICE

The data show that present practice is purely a matter of chance. Solid line segments, dotted line segments, parentheses, no provision of any kind, letter symbols, squares and rectangles of various magnitudes, constitute the kinds of items in practice. These are placed before, after, and below the statement. It is even possible to conceive of a placement above the statement. The practice shows the provisions in vertical staggered columns, and at any chance distance before the first or after the last letter in a statement.

### STATEMENT OF PURPOSES

The purposes of this chapter are:

1. To determine an adequate response provision for the +0 written response.[1]

### DETERMINATION OF THE RESPONSE PROVISION

*Procedure*—1. A random sample of 200 responses was taken from an examination in the subject matter course in Columbia College called "Contemporary Civilization."

2. The distance of each response from the left-hand end of the response space dotted line to the point of intersection of the horizontal and vertical marks of the plus sign and the major and minor axes drawn within the zero sign were measured. The response space dotted line was 25/64 inch in length.

3. The measurements were made in sixty-fourth parts of an inch.

[1] Determination of the response provision for the t-f printed response is discussed on pages 35-36 of manuscript copy of this dissertation, Teachers College Library.

*Result.*—The distribution of these measurements was:

1. Arithmetic Mean was 15/64 inch from left-hand end of response space.
2. Range of plus and minus three times Standard Deviation equaled ¼ inch.

| 64ths inch | Frequency |
|---|---|
| 7 | 1 |
| 8 | 2 |
| 9 | 5 |
| 10 | 12 |
| 11 | 14 |
| 12 | 27 |
| 13 | 13 |
| 14 | 20 |
| 15 | 22 |
| 16 | 28 |
| 17 | 17 |
| 18 | 21 |
| 19 | 6 |
| 20 | 4 |
| 21 | 3 |
| 22 | 2 |
| 23 | 2 |
| 24 | 1 |
| Total | 200 |

*Conclusion.*—The length of a response space, sufficiently long to include with certainty the chance deviations of a student while writing his response, is 5/16 inch.

TYPEWRITER ADAPTATIONS

1. The +0 response dotted line segment

.... 69. **********
**********
********

22223452333

2. The t-f printed response

t f 9. **********
**********

1333363342333

1 = Typewriter key upon which appears the letter *t*.
6 = Typewriter key upon which appears the letter *f*.
3 = Typewriter space bar.
4, 5 = Typewriter key or keys upon which appear the digits desired.
2 = Typewriter key upon which appears the period.

Note:—No measurement is given for the number of the statement. The number of digits in a statement varies from 1 to 3, inclusive.

PLACEMENT OF THE RESPONSE PROVISION

*Procedure*—1. The judgments of the five expert judges were obtained.

2. Analyses were made of the movements of a scoring key by a right-handed scorer while scoring, with the response provisions before and after a series of true-false statements.[2]

*Results.*—1. Provision for the response:

Before the statement———3 judges.

After the statement———2 judges.

CHART 2—ANALYSIS A

PROVISIONS FOR THE INDICATIONS OF THE RESPONSE BEFORE THE STATEMENT IN A VERTICAL COLUMN

| Number of the Motion | MOTION OF THE | |
|---|---|---|
| | RIGHT HAND | LEFT HAND |
| 1 | *a.* 12 inches; upper end to lower end of page 1, parallel to scoring key, to the right and adjacent to examiner's indications of responses. | *b.* No motion. |
| 2 | *a.* 12 inches; lower end to upper end of page 1, parallel to scoring key, to the right and adjacent to the examiner's indications of responses. | *b.* No motion. |
| 3 | *a.* 11 inches; upper end to lower end of page 1; where the thumb moves (1 inch) to grasp page 1 preparatory to turning it. | *b.* 1½ inches as the minimum of the scoring key from its *proper position* to the left edge of the length of the examination paper. |
| 4 | *a.* Approximately 25-inch diagonal, lower left-hand corner to upper right-hand corner and extended in a straight line approximately ten inches. | *b.* 1½ inches motion, as the minimum of the scoring key from the left edge of the length of the examination paper to the *proper position* of the scoring on page 2. |
| 5 | *a.* 14 inches; from the end point of the line which had been extended 10 inches to the top of the page to the right and adjacent to the examiner's first indication of responses on page 2. | *b.* No motion. |

[2] Details of the motion analysis procedures are given in full on pages 38, 245-57 of manuscript copy of this dissertation, Teachers College Library.

2. The motion analysis indicated one essential motion difference of approximately eleven inches for each page of the examination in favor of placement of the provision of the response before the statements for a right-handed scorer.[3] (See Charts 2 and 3.)

CHART 3—ANALYSIS B

PROVISIONS FOR THE INDICATIONS OF THE RESPONSE AFTER THE STATEMENT IN A VERTICAL COLUMN

| Number of the Motion | MOTION OF THE | |
|---|---|---|
| | RIGHT HAND | LEFT HAND |
| 1 | *a.* Same as motion 1*a* in Analysis A. | *b.* Same as motion 1*b* in Analysis A. |
| 2 | *a.* Same as motion 2*a* in Analysis A. | *b.* Same as motion 2*b* in Analysis A. |
| 3 | *a.* Same as motion 3*a* in Analysis A. | *b.* 7 inches, as the minimum of the scoring key from its *proper position* to the left edge of the length of the examination paper. |
| 4 | *a.* Same as motion 4*a* in Analysis A. | *b.* 7 inches, as the minimum of the scoring key from the left edge of the length of the examination paper to the *proper position* of the scoring key on page 2. |
| 5 | *a.* Same as motion 5*a* in Analysis A. | *b.* No motion. |

TIME AND LABOR FACTORS IN ASSEMBLING STATEMENTS

When the response provisions are placed after the statements in a non-staggered column, it is the practice in many instances to continue from the period which ends the statement to the provision for the response, by means of a series of hyphens, dots, or dashes. This process conspicuously relates the provision for response to the statement for which it was intended.

Example: Metternich's fall from power came with the tidings of the French Revolution of 1830,.....

[3] The per cent of right-handed children is more than 90 per cent compared to less than 10 per cent of left-handed children.
Gould, George M. *Right-Handedness and Left-Handedness*, p. 210.
"All that is needed to explain right-handedness in 94 per cent of children . . ."

When this operation is repeated in terms of 150 to 200 statements for an examination manuscript, a small amount of time and labor is obviously consumed in the process of typing hyphens.

When the response provision is placed before the statement no hyphens are necessary. Granting that hyphens require time and labor, in addition to that required to type the statement and the response provision, then the placement of the provision for the indication of the response would be placed before the statement.

*Recommendation.*—It is recommended that the provision for the indication of the response be placed *before* the statement in a non-staggered column.

If additional research should show that the above recommendation does a real and significant wrong to the accuracy of placement of the student's indications of response, then the economics of the situation probably should yield to the behavioristic aspects of the problem.

# CHAPTER IV

## THE ASSEMBLING OF STATEMENTS

### PRESENT PRACTICE

The variability of assembly is found in the following factors:

1. Indention of the statement—paragraph, no indention, hanging indention.
2. Placement of the Provision for response—before, after, below, the statement.
3. Placement of the number of the statement.
   *a.* Before or after the statement.
   *b.* Before or after the Provision for response.

In Chapter III the Provision for response was arbitrarily placed before the statement.

### STATEMENT OF PURPOSE

The purpose of this chapter is to attempt to find an acceptable method for assembling statements.

*Procedure.*—1. The materials for the experiment were assembled into three Packets. Packet No. 1 contained six 5 x 8 inch cards, and Packets 2 and 3 each contained twelve 5 x 8 inch cards. Each card contained one sample of a method of assembling statements.[1] (See Appendix D.)

2. A carefully planned memorandum to the expert judges was prepared. The memorandum specified the procedure for judgment of the materials in the Packets.[2]

*Results.*—The expert judges were unable to say explicitly why one form of assembly was preferred to another. It apparently depended in part upon the type of assembly that they were accustomed to using. Two judges admitted that the general æsthetic appearance of the page seemed to influence their judgment. These, no doubt, with many other subtle factors, played

[1] Complete list of materials is given in the manuscript copy of this dissertation, Teachers College Library.
[2] Details of procedure are explained in manuscript copy, pages 41-43, 248-49.

a part in addition to the tentative criteria of information provided in the experiment. The summary of the judgments of the four expert judges and the writer, is given in Chart 4 by Operation units of the experiment, and Card Number and Samples.

CHART 4

CARD NUMBER AND SAMPLE OF ASSEMBLY FORM FOR A TRUE-FALSE STATEMENT, SELECTED FROM EACH OPERATION UNIT OF THE EXPERIMENT, BY EACH OF FIVE EXPERT JUDGES.

| Operation No. | | Number of the Judge | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5† |
| 1..... | Card No. | 1293 | 1293 | 1293 | 1992 | 3129 |
| | Sample | 1. *** * * * | | | 1. **** **** | 1. ** ***** |
| 2..... | Card No. | 9991 | 9991 | 9991 | 9991 | 2199 |
| | Sample | .... 1. * *** ** | | | | 1. .... ** ********** |
| 3..... | Card No. | 1267 | 7843 | 7843 | 7843 | 1679 |
| | Sample | t f *** * | t f 1. | | **** ** | 1. t f ** ********** |
| 4..... | Card No. | 9991 | 9991 | 9991 | 7843 | 2199 |
| | Sample | .... | 1. **** ** | | t f 1. **** ** | 1. .... ** ********** |

† Very difficult for this judge to do this experiment.

CARD NO. 9991 FROM PACKET NO. 2

.... 1. ********** ********** .
.... 2. ********** ********** ********** .
†.... 3. ********** .
†.... 4. ********** ********** ********** .
.... 5. ********** ********** ********** ********** .
.... 6. ********** ********** .
... 7. ********** ********** ********** .

† For convenience of scoring avoid putting two response spaces nearer to each other than one and one-half typewriter line spaces.

Card No. 9991, . . . . I. *****, was selected as the best by three judges. ***

*Conclusions.*—The results of the judgments of the five experts indicate that:

1. The assembly on Card No. 9991, p. 24, is preferred.
2. All the judges agree that each statement should have a number.
3. Four judges agree that the number should follow the dotted line.
4. Four judges agree to the use of two space hanging indention of all lines following the first line of a statement.

# CHAPTER V

## SPEED AND ACCURACY OF SCORING

### INTRODUCTION

In Chapter II two distinct methods were selected through judgment to be the best by which to indicate a response to a true-false statement. The problem still remains which method is quicker and more accurate to score. Either the Freehand Written responses, + and 0, or the Mechanical responses of t-f may be scored without the use of a mechanical scoring key. Ben D. Wood has collected evidence to show that a mechanical scoring key is two or more times as fast and as accurate as scoring without such a key.

When mechanical keys are used the responses are placed in a non-staggered vertical column. The eye may travel more nearly in a straight line down the page. This practice should tend to minimize the use of the eye muscles which would control movements of the eye to any other part of the examination page.

### STATEMENT OF PURPOSE

The purpose of this chapter is to determine the relative speed and accuracy of scoring the responses to true-false statements by the +0 and t-f scoring methods.

### RELATIVE SPEED OF THE +0 AND t-f SCORING METHODS

*Procedure.*—1. Fifty examination papers, to be scored by the +0 key and the t-f key, respectively, were constructed with a specified procedure. (See sample, Appendix E.) Each examination paper consisted of 155 response provisions identically placed on each of the five pages.[1]

2. Chapter VI describes how to construct the +0 key.[2]

[1] This procedure is discussed in detail in the manuscript copy of this dissertation, on file at Teachers College Library. Pages 256-58, 261.

[2] Description of how to construct the t-f key, manuscript copy, pages 277-82.

3. Each group of fifty papers was divided into two sub-groups of twenty-five papers each, so that the scoring process might be rotated.[3]

4. A supply of pencils was provided.

5. Three tested stop watches were selected.

6. Care was exercised in seating the scorer and placing the materials for scoring.

7. The three timers were each given an opportunity to practice their respective tasks before the scoring experiments were performed. The timers had no difficulty in checking to within one-half of one second with each other. All times in the experiment were to the nearest whole second.

8. The two scorers were each given the opportunity of practising the process of scoring until each was ready to begin the experiments.

9. The timers and scorers also were given the opportunity of practising together in a preliminary experiment.[4]

10. A total of seven definable parts of the scoring process were timed by the three timers. Each timer recorded but one unit during the scoring process. The remaining four were obtained through processes of addition or subtraction. (See Appendix E., Definition of Timing Units.)[5]

11. The watches were not stopped during the experiment. The times of the watches at particular points in the total job of scoring a particular paper were recorded.

12. Critical Ratio technique was applied to the time data.[6]

*Results.*—1. Charts 5 and 6 show the relative distribution of time during the process of scoring by the +0 key and t-f key, respectively.

2. Table 14 shows the significant time differences between the two key methods of scoring for each of the Time Units.[7]

[3] McCall, Wm. A. *How to Experiment in Education,* Chapter II.

[4] Steps 5, 6, 7, 8, and 9 are explained in full detail on pages 262-67 of manuscript copy of this dissertation, Teachers College Library.

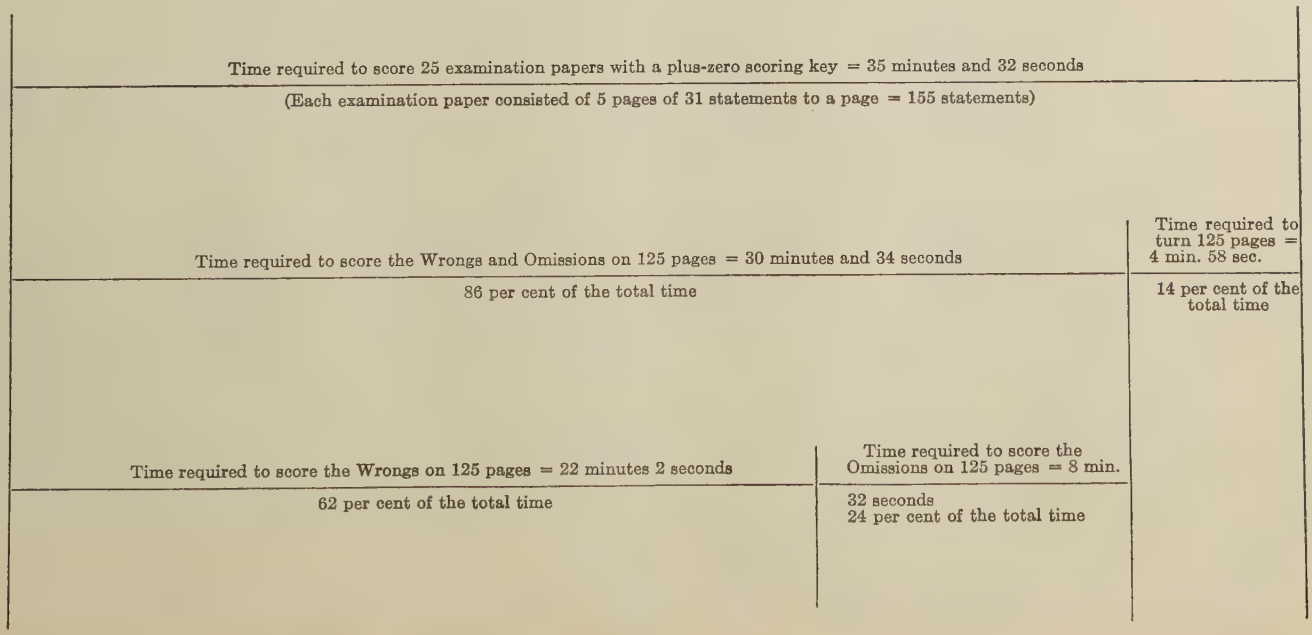
[5] Mathematical process of obtaining the time data in each of the seven Timing Units is explained on pages 271-75 of the manuscript copy.

[6] "The Critical Ratio is the quotient obtained by dividing the difference between the averages on an item by the Probabe Error of that difference." McGaughy, J. R. The Fiscal Administration of City School Systems, page 55.

[7] The basis for Table 14 is a series set forth on pages 58-64 of the manuscript copy of this dissertation, Teachers College Library.

## CHART 5

PER CENT DISTRIBUTION OF THE ROTATED TIME BY TIME UNITS REQUIRED TO EXECUTE THE COMPLETE PROCESS OF SCORING 25 TRUE-FALSE EXAMINATIONS WITH A PLUS-ZERO SCORING KEY



## CHART 6

PER CENT DISTRIBUTION OF THE ROTATED TIME BY TIME UNITS REQUIRED TO CARRY ON THE COMPLETE PROCESS OF SCORING 25 TRUE-FALSE EXAMINATIONS WITH A T-F SCORING KEY

Time required to score 25 examination papers with a t-f scoring key = 29 minutes and 33 seconds

(Each examination paper consisted of 5 pages of 31 statements to a page = 155 statements)

| Time required to score the Wrongs and Omissions on 125 pages = 26 minutes and 2 seconds | | Time required to turn 125 pages = 3 minutes 32 seconds. |
|---|---|---|
| 88.1 per cent of the total time | | 11.9 per cent of the total time |
| Time required to score Wrongs on 125 pages = 14 minutes 33 seconds | Time required to score the Omissions on 125 pages = 11 minutes and 29 seconds | |
| 49.3 per cent of the total time | 38.8 per cent of the total time | |

*Conclusion.*—The t-f key is faster in every part of the scoring process, except that of scoring *omissions.* In every case the Critical Ratios are very significantly in favor of either the t-f or the +0 key. (See Table 14.)

TABLE 14

SIGNIFICANT TIME DIFFERENCES BETWEEN THE TWO KEY METHODS OF SCORING FOR EACH OF THE TIME UNITS

| No. | Time Unit | Scoring Key Having Larger Average | Scoring Key Having Smaller Average | Critical Ratio of Difference Between Averages |
|---|---|---|---|---|
| 3..... | Page to the next consecutive page | +0 | t-f | 16.8 |
| 4..... | Scoring Wrongs and Omissions | +0 | t-f | 8.4 |
| 5..... | Scoring Wrongs | +0 | t-f | 22.4 |
| 6..... | Scoring Omissions | t-f | +0 | 15.5 |
| 7..... | Turning a Page | +0 | t-f | 9.9 |
| 5..... | Special Case—scoring from zero to two Wrongs inclusive per page | +0 | t-f | 17.0 |
| 6..... | Special Case—scoring zero Omissions per page | t-f | +0 | 13.8 |

RELATIVE SPEED OF SCORING A WRONG AND AN OMISSION WHEN THE DECISION NECESSITATES RECOGNIZING ONE SYMBOL

When a *wrong* is marked by the t-f key scoring method, only the t's and f's which have been *wrongly* encircled show through the circular punched hole of the key. The scorer looks only for a circle irrespective of t's or f's.

The +0 key necessitates recognition of two symbols in order to score a *wrong.*

When an *omission* is marked by the +0 scoring method, the scorer is required to recognize only a response provision with neither a + nor a 0 written upon it.

The t-f key necessitates recognition of two symbols in order to score an *omission.*

*Results.*—1. Table 14, Time Unit No. 5, Critical Ratio 22.4, conclusively favors the t-f key scoring method.

2. Table 14, Time Unit No. 6, Critical Ratio 15.5, conclusively favors the +0 key scoring method.

*Conclusion.*—Scoring a *wrong* and an *omission* is significantly faster when the decision is based upon the recognition of one symbol.

### Relative Accuracy of the +0 and t-f Scoring Methods

*Procedure.*—The fifty examination papers scored by the +0 key and the t-f key, respectively, were checked for errors of scoring.

*Results.*—1. For each scoring key the number of responses was 7,750; the number of *wrongs* was 1,752; and the number of *omissions* was 372.

2. The dominant types of errors of scoring by each key were
   *a.* Unmarked *omissions.*
   *b.* Unmarked *wrongs.*

3. Table 16 below shows the number of perfectly scored examination papers out of 50 by the +0 key and the t-f key scoring methods, on several bases.

TABLE 16

| Number of examinations out of 50 that were scored perfectly | Scoring Key | |
|---|---|---|
| | +0 | t-f |
| In every respect | 38 | 17 |
| With respect to unmarked *omissions* | 47 | 20 |
| With respect to unmarked *wrongs* | 39 | 44 |

4. Table 17 displays the distributions of unmarked *omissions* per examination by the +0 key and the t-f key scoring methods, for fifty examination papers.

5. Table 18 displays the distribution of unmarked *wrongs* per examination by the +0 key and the t-f key scoring methods, for fifty examinations.

TABLE 17

| Number of Errors Unmarked Omissions | Frequency of Unmarked Omissions per Examination | |
|---|---|---|
| | +0 key | t-f key |
| 0 | 47 | 20 |
| 1 | 2 | 18 |
| 2 | 1 | 6 |
| 3 | .... | 5 |
| 4 | .... | 0 |
| 5 | .... | 1 |
| Total | 50 | 50 |
| Mean | 0.08 | 1.00 |
| Standard Deviation | 0.34 | 1.11 |
| Critical Ratio=C.R.=8.3 | ..... | ..... |

TABLE 18

| Number of Errors Unmarked Wrongs | Frequency of Unmarked Wrongs per Examination | |
|---|---|---|
| | +0 key | t-f key |
| 0 | 39 | 44 |
| 1 | 9 | 5 |
| 2 | 1 | 1 |
| 3 | 1 | 0 |
| Total | 50 | 50 |
| Standard Deviation | 0.28 | 0.14 |
| Critical Ratio=C.R.=2 | | |

*Conclusions.*—1. In Table 17 the Mean, Standard Deviations, and Critical Ratio are decidedly in favor of the +0 key. It seems significant to say that the +0 key method of scoring leaves unmarked fewer *omissions* than does the t-f key scoring method.

2. In Table 18 The Critical Ratio (See Table 14) is less than 3. There is some significance to the fact that the t-f key scoring method tends to leave unmarked fewer *wrongs* than does the +0 key scoring method.

### THE RELATIVE ACCURACY OF SCORING A WRONG AND AN OMISSION WHEN THE DECISION NECESSITATES RECOGNIZING ONE SYMBOL[8]

*Procedure.*—No new procedure was required.

*Results.*—1. Table 16 favors the +0 key for scoring examination papers perfectly in every respect.

2. Table 17 shows a total of
   4 unmarked *omissions* (one symbol) by the +0 key.
   50 unmarked *omissions* (two symbols) by the t-f key.
3. Table 18 shows a total of
   7 unmarked *wrongs* (one symbol) by the t-f key.
   14 unmarked *wrongs* (two symbols) by the +0 key.

*Conclusion.*—Scoring a *wrong* and an *omission* is more accurately performed when the decision is based upon the recognition of one symbol.

### OPINIONS OF THE TIMERS AND ONE SCORER

*Procedure.*—An additional interesting check based on observation extended over one experiment with the +0 key and one experiment with the t-f key was obtained.

Immediately after the two above experiments were concluded each of the three timers and the scorer were asked to record their opinions upon a sheet of paper. Their opinions concerned speed and accuracy of scoring by the +0 key scoring method compared to the t-f key scoring method. None of the experimental evidence had been handled statistically as yet. It must be remembered, however, that the timers recorded the primary data for speed of scoring.

There was, apparently, no direct way to judge accuracy.

*Summary of results.*—The opinions of the three timers and the scorer concerning the t-f key were as follows:

Timer No. 1:
   *a.* Less strain scoring the *wrongs.*
   *b.* Some strain in checking *omissions.*

Timer No. 2:
   *a.* t-f key relieved the scorer of strain.

Timer No. 3:
   *a.* t-f key easier to score *wrongs.*

[8] See page 31 for one-symbol definition.

*b.* t-f key apparently required less mental effort when properly adjusted.

*c.* More movement required to see *omissions.*

The opinions of three timers on the +0 key were:

Timer No. 1:

*a.* More strain on the scorer.

Timer No. 2:

*a.* This key seems to tire the scorer.

Timer No. 3:

*a.* Easier to adjust.

*b.* Less movement required to see the *omissions.*

*c.* Key easier to handle.

## CONCLUSIONS

1. The t-f key scoring method is undoubtedly very much faster than the +0 key scoring method when scoring *wrongs.*

2. The t-f key scoring method is undoubtedly very much slower than the +0 key scoring method when scoring *omissions.*

3. The t-f key is not significantly more accurate than the +0 key with respect to scoring *wrongs.*

4. The t-f key is very significantly less accurate than the +0 key with respect to scoring *omissions.*

5. It would seem defensible to say that significant accuracy is preferred to significant speed in the scoring of any kind of an examination from the standpoint of the satisfaction of the student as well as the instructor. Therefore the +0 key is to be preferred to the t-f key in scoring true-false statements, until a further research answers the question "How fast is the +0 key compared to the t-f key when accuracy is rendered the same for both of the above scoring methods?"

6. The opinions of each of the three timers and the scorer support the above conclusions in so far as opinions may be used as a real support for anything.

7. Increase in speed is accompanied by an increase in accuracy when the process of scoring is based upon recognition of one symbol for either a *wrong* or an *omission.* (See Tables 14, 17, and 18.)

# CHAPTER VI

## PREPARATION OF SCORING KEYS

### PLUS-ZERO SCORING KEY

List of materials:

1. A 12-inch scaled straight-edge.
2. A lead pencil.
3. A pair of scissors.
4. Three pieces of common twine, each 6 inches in length.
5. Pen and ink (black or blue) or fountain pen.
6. A strip of light cardboard for each edge of the examination paper. Each strip is one inch wide. Their length will depend upon the length of the pages of the examination paper. The cardboard is about as heavy as manila folder paper.
7. One strip of light cardboard, one inch wide and eight inches long.
8. Penknife. If the scissors have one pointed blade the knife is unnecessary.
9. Red or blue pencil.

### HOW TO CONSTRUCT THE PLUS-ZERO KEY

1. Chart No. 7 shows the outline of the plus-zero scoring key.
2. The material used to make the key is light cardboard about the thickness of the cardboard of a manila folder.
3. Each page of the key is one inch wide and the same length. The length is the distance between the scoring key line at the upper end of each examination paper page and the lower edge of the page.

The scoring key line is mimeographed or typewritten upon each page of the examination. Its position should be as nearly the same as possible upon each page of the examination.

4. The examination page number is placed at the mid-point near the top edge of each page of the key. The number one is

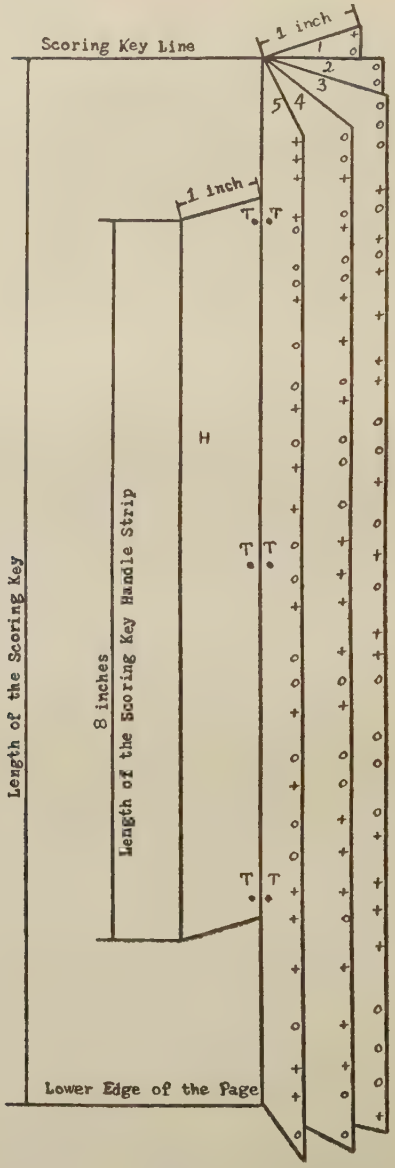



CHART 7. PLUS-ZERO SCORING KEY

placed on the last page of the key. The largest number is placed upon the top facing page of the key.

5. The pages of the key are stacked together and firmly held. Three holes, T, are cut with a knife blade or scissor blade point at the upper and lower ends, and mid-point of the pages of the key. These holes are cut near the edge of the key opposite the scoring key symbols, namely plus's and zero's. (See Chart No. 7.)

6. The handle for grasping the key is 1 inch by 8 inches. Three holes are cut through it corresponding to the three holes cut through the pages of the key.

7. The pages of the key and handle are bound together with the twine. Pass a 6-inch piece of twine through each of the three holes, and tie a square knot in each piece of twine. The pages should spread easily.

### HOW TO MARK THE KEY WITH PLUS AND ZERO SIGNS CORRESPONDING TO THE RIGHT ANSWER TO EACH STATEMENT

1. Take a blank examination paper and put the correct answer on the dotted line before each statement. This provides a correctly answered examination. It is called the *examination key*.

2. *Be sure* that each page of the examination key is numbered above the scoring key line. It is well also to make sure that every examination paper given to students has been checked with respect to correct page number order. This needs to be done in order to expedite scoring, no matter what kind of a scoring method is used.

3. Place the scoring key page No. 1, upon the examination key page No. 1. The upper edge of the scoring key should coincide with the scoring key line. The scoring key should lie adjacent, to the left, and parallel to, the non-staggered column of 5/16 inch dotted lines (response provisions).

4. With black or blue ink proceed to write the plus and zero signs of the examination key upon the scoring key as follows:

*a.* All plus's and zero's are placed in a vertical column, adjacent to the right-hand edge of the scoring key.

*b.* Every plus and zero of the examination key is copied so that its position on the scoring key is as nearly oppo-

site to its position on the examination key as is possible. When the copying is finished, scoring key page No. 1 will appear exactly like examination key page No. 1, i.e., plus for plus and zero for zero.

5. Turn to page 2 of the scoring key.
6. Turn to page 2 of the examination key.
7. Repeat the above procedure of placing the scoring key and copying the plus's and zero's. When the copying is finished scoring key page No. 2 will appear exactly like examination key page No. 2.
8. The same procedure is followed for the remainder of the corresponding scoring key and examination key pages.

The result is a complete plus-zero scoring key which is ready to use.

### HOW TO USE THE PLUS-ZERO SCORING KEY

1. Place page No. 1 of the scoring key as in Step 3, on page No. 1 of the examination paper handed in by a student.

2. Every time a plus sign on the scoring key is opposite a zero on the examination paper, the statement was incorrectly answered by the student. Also, every time a zero on the scoring key is opposite a plus on the examination paper, the statement was incorrectly answered by the student. The above are called *wrong* answers.

All *wrongly* (incorrectly) answered statements are indicated by the scorer with a straight line segment about one-half inch to one inch in length, made with a red or blue pencil. The line segment is placed to the right of the scoring key.

3. Every case in which the student failed to put either a plus or a zero upon the dotted line of a statement, is considered as an *omission.*

All *omissions* are indicated by the scorer with a zero, made with a red or blue pencil. The zero is placed to the right of the scoring key.

4. All statements that have neither a red or a blue line segment nor a red or a blue zero opposite them are correctly answered statements. Such statements are called *right* answers.

### T-F SCORING KEY [1]

The experimental evidence indicates that the t-f scoring key is significantly less useful for the purpose of scoring classroom tests than the +0 key.

### RELATIVE DIFFICULTY OF CONSTRUCTING THE PLUS-ZERO SCORING KEY AND THE T-F SCORING KEY [2]

*Materials.*—1. Less materials are required to construct the +0 key.

2. The materials needed to construct the plus-zero key are available among personal supplies, without added expense. The t-f key requires the probable purchase of a circular ⅜ inch hollow punch, a small hammer, and a 2″ x 10″ x 15″ block of hard pine. In either case the light manila-weight cardboard would probably be purchased.

3. The amount of accuracy required within the limits of personal error is less for the plus-zero key than for the t-f key.

*a.* The plus-zero key requires accurate proper position of the key and accurate copying of plus's and zero's to corresponding positions opposite the copy.

*b.* The t-f key requires accurate drawing of parallel lines, accurate circling of t's and f's, accurate superposition of the pages of the examination key upon the corresponding page positions of the scoring key, and accurate punching of holes.

4. The amount of labor required to assemble materials and construct the +0 key is less than for the t-f key.

5. The different kinds of skills required for the process of construction of the +0 key are less than required for the t-f key.

The above five comments are based largely upon the experience of the writer. It is partially a result of an analysis of the steps in the job of constructing each key, and partially the result of the opinions of experts.

It seems defensible to say that the +0 key is easier to construct because of the facts (1) that materials needed are usually at hand and (2) the skills required are those probably possessed by most people.

[1] Details of preparation and use of this key are explained on pages 276-83 of the manuscript copy of this dissertation on file at Teachers College Library.

HOW TO COMPUTE THE STUDENT'S SCORE FOR EITHER THE PLUS-ZERO OR T-F SCORING KEYS

1. Add the Rights = R
2. Add the Wrongs = W
3. Score = Total of the Rights minus Total of the Wrongs.[2]

The formula is S = R — W.

[2] Wood, Ben D. "Studies of Achievement Tests." *Journal of Educational Psychology*, pp. 1-22.

# PART TWO
# CONTENT

## CHAPTER VII

### MECHANICS OF COMPOSITION OF STATEMENTS

#### PURPOSE

The purpose of this chapter is to justify the use of "rules" by which to compose true-false statements.

#### PROCEDURE

The body of this chapter was assembled in the following manner:

Handbooks on English composition [1] were consulted for the purpose of searching out principles or "rules" which represent good and bad practices in English composition.

The true-false statement data were used for possible illustrations of each practice.

Wherever a true-false illustration of a given practice was found, a brief rule was formulated.

The concrete true-false statement illustration was written below the rules in both the "given" and an "improved" form.

It seemed unnecessary to attempt to present a complete job analysis in this way in this study. No claim is made for a complete set of rules. The procedure, however, serves to point out the close relationship which good English composition bears to the formulation of true-false examination statements.

[1] Royster, James F. and Thompson, Stith. *Manual and Handbook for English Composition.*
University of Iowa Extension Bulletin No. 68, University of Iowa.
*The High School Journal,* "Minimum Essentials of English Composition," Vol. 5, Illinois Association of Teachers of English, Vol. XI, Bull. No. 1.
Standard Usage in English by Department of English, University High School, University of Chicago Press.

CLASSIFICATION OF RULES

The rules are classified in groups under the headings: (1) Punctuation, (2) Spelling, (3) Sentence structure, (4) Grammar, (5) Diction.

### *Punctuation*

Five sets of true-false examinations were used to determine what punctuation marks were used. Four of these were reported by Ben D. Wood[2] as having validity coefficients as follows:

| | Validity Coefficient |
|---|---|
| Pleading and practice 180 items | +0.87 |
| Property test ..... 200 items | +0.89 |
| Torts test ..... 130 items | +0.82 |
| Anatomy test ..... 130 items | +0.78 |

The additional test used was an examination in equity of 180 items. It is assumed that the above examinations are true-false examinations of a superior quality.

The analysis of the distribution of punctuation marks in the five true-false examinations of assumed superior quality is shown in Table 21.

Evidence presented indicates that as far as need be, the comma, apostrophe, and quotation marks are inserted as follows:

TABLE 21

DISTRIBUTION OF PUNCTUATION MARKS IN EQUITY, TORTS, PLEADING AND PRACTICE, REAL PROPERTY, AND ANATOMY TRUE-FALSE EXAMINATIONS

| True-False Examination in | Comma | Apostrophe | Quotation Marks | Semicolon | Colon | Dash | Asterisk |
|---|---|---|---|---|---|---|---|
| Equity | 74 | 22 | 27 | 6 | 2 | 0 | 0 |
| Torts | 76 | 13 | 2 | 0 | 0 | 0 | 0 |
| Pleading and Practice | 161 | 52 | 17 | 5 | 2 | 0 | 2 |
| Real Property | 142 | 36 | 4 | 4 | 2 | 0 | 0 |
| Anatomy | 17 | 0 | 0 | 0 | 3 | 1 | 0 |
| Total | 470 | 123 | 50 | 15 | 9 | 1 | 2 |
| Per Cent | 70.2 | 18.4 | 7.5 | 2.2 | 1.4 | 0.1 | 0.2 |

[2] Wood, Ben D. "Studies in Achievement Tests," *The Journal of Educational Psychology* Vol. XVII, No. 1.

*a.* Use the comma to clarify the thought in a sentence.
*b.* Use the apostrophe to denote possession, or omission of letters.
*c.* Use the quotation marks in direct quotation.
*d.* Avoid complex and long sentence structures as far as possible. This will tend to eliminate unusual punctuation.
*e.* Use capital letters for
(1) Names of places, people, months, days of the week, holidays, beginning of principal words of titles of books and printed articles, and proper nouns.
(2) First word of a sentence.
(3) First word of a direct quotation.
(4) Degrees, titles, and initials.
(5) Important words in topics of papers.
(6) First word of resolutions, questions, etc.
(7) School subjects when derived from proper nouns.

*Spelling*

1. Avoid misspelled words.

False: The Binet-Simon texts are the original and best-known.
True: The Binet Simon tests are the original and best known tests.

The similarity of the "s" and "x" in script leads to a typographical error in spelling.

2. Avoid an error in homonyms.

False: A city superintendent of schools should be an affective leader of public education in the community.
True: A city superintendent of schools should be an effective leader of public education in the community.

3. Avoid abbreviations.

Some exceptions to this rule are: i.e., e.g., viz., etc., A.D., B.C., Mr., Mrs., Messrs., Dr., Rev., Hon., Esq., St., Mgr., and the French M., Mme., Mlle.

False: Congress meets in Washington, N. C.
True: Congress meets in Washington, D. C.

The instructor wrote D. C. and the typist misread the D for an N.

4. Be sure the division of a word at the end of a line is indicated by a hyphen and that the division is according to the syllable.

5. Avoid the division of a word at the end of a line as far as possible.

True: The first draft of a treaty is usually in inacceptable form.

The typewritten copy read: "The first draft of a treaty is usually in inacceptable form." The hyphen after the syllable "in" was omitted when the word was broken at the end of the line, and the proof reader crossed out the syllable "in" because it seemed that the word "in" had been repeated. The statement then read:

False: The first draft of a treaty is usually in acceptable form.

Another example of this word division resulted as follows:

True: The child of average intelligence is a typical child of a group of children. In the final form in the examination it read:

False: The child of average intelligence is an atypical child of a group of children.

6. Avoid "Tricky" use of the singular and plural of words.

False: The Corporation Acts of 1661 provided for the restriction of commercial corporations in England.

True: The Corporation Act of 1661 provided for the restriction of commercial corporations in England.

### *Sentence Structure*

7. Avoid loose sentence structure. A sentence may contain two complete ideas. The unity of thought is better acquired and the sense more complete by the use of two complete statements.

False: Multiplication of decimals should be taught without reference to whole numbers and should be taught inductively.

False: (Revised) Multiplication of decimals should be taught without reference to whole numbers.

True: Multiplication of decimals should be taught inductively.

8. Avoid all modifiers which do not aid to clarify the thought expressed in the statement.

False: Man's long period of infancy makes possible a larger number of reaction abilities to meet the enormous and perplexing number of the demands of life.

Revised: Man's period of infancy makes possible a higher degree of development to meet the complex demands of life.

9. Put a complete thought in each statement.

False: Chapman and Counts regard the "complex" as distinctly pathological, and insist that it has a place in the life of the normal individual.

False: (Revised) Chapman and Counts regard the "complex" as distinctly pathological.

True: Chapman and Counts insist that the "complex" has a place in the life of the normal individual.

10. Do not string together more than two clauses in a compound sentence.

Given: One measure of central tendency is the mean, another is the median, and a third common measure of central tendency is the mode.

Improved: Mean, median, and mode are the three commonly used statistical measures of central tendency.

If any of the above clauses were false, the statement would be scored as false. An improved structure might be:

The three commonly used statistical measures of central tendency are:

The Mean
The Mode
The Median

11. When short statements are closely related, put them into one statement.

Given: A poor reasoner often takes unfruitful clues.
A poor reasoner does not check results.
A poor reasoner is not critical.

Improved: A poor reasoner takes unfruitful clues, does not check results, and is not critical.

12. Seek coherence in the statement by avoiding the use of the indefinite pronoun "it."

Given: A negative correlation indicates that there is no correlation to it.

The above statement is poor for many reasons:

*a.* Indefinite "it."
*b.* The type of correlation is not specified.
*c.* The statement does not state the assumption "other things being equal."
*d.* The expression "that there" is not good usage.
*e.* In its present form the statement is so ambiguous that it is doubtful whether it is true or false. It probably was intended to be false. Most persons probably would mark it false.

Improved: Other things being equal, a negative value of the Pearson Product-Moment Correlation Coefficient between two factors indicates no correlation.

Another example of poor coherence in the statement because of the use of the indefinite pronoun "it":

Given: It has been found that 80 per cent of the school children test between 90 and 110 I. Q.

Improved: According to test results about 80 per cent of school children have I. Q.'s between 90 and 110.

13. Avoid the ambiguous use of "that" or "those."

Given: The Kaiser was one of those typical Prussian Monarchs.

The statement would lead one to believe there was a group of typical Prussian Monarchs. The total number of Prussian Monarchs is too small to make even a small number of typical ones. If the statement is not false, then it is ambiguous. It was probably intended to be stated:

Improved: The Kaiser was a typical Prussian Monarch.

The statement is still somewhat doubtful. Historical opinion may agree sufficiently to warrant a response of true. If the opinion of the instructor was in agreement with the above statement, then it could be true on that basis.

14. Avoid ambiguous reference of pronouns to antecedents.

Given: Even Bismarck never dared to defy the elected representatives of the people although very reactionary.

Was "Bismarck" or the "representatives" reactionary?

Improved: Although "Bismarck" was very reactionary, he never dared to defy the elected representatives of the people.

15. Avoid dangling clauses.

Given: State Board members should be paid a salary which is a living wage, if elected by the people.

Improved: State Board members should be paid a salary which is a living wage.

15*a*. Avoid vague, dangling participles,—a participle which refers to an idea conveyed by a sentence as a whole or implied by some word in the sentence. Make it clear who was meditating.

Given: The chief duty of the principal of a school is planning for the improvement of his supervision.

Does it mean supervision of himself or school, or is supervision of instruction implied?

Improved: The chief duty of the principal of a school is improvement of supervision of instruction.

16. Avoid an omitted subject or predicate.

Given: . . . 26. The retentiveness of memory is fixed by heredity and is unchanged by training.

. . . 27. The same for intelligence.

No. 27 is not a statement. It is ambiguously related to statement No. 26. Does it mean the "retentiveness of intelligence" or "fixed by intelligence" or "unchanged by intelligence"? No. 27 seems to be more a test of mental gymnastics upon a haloed bar of ambiguous doubtfulness than a direct test of subject matter.

17. Avoid an abbreviation of names of titles of books or other publications. Such practice may be interpreted as disrespectful to the author or his publication.

Given (Ambiguous): Kelley published a book entitled "Statistics."

Which Kelley published the book?

False: Truman Lee Kelley published a book entitled "Statistics."

True: Macmillan published a book entitled "Statistical Method" written by Truman Lee Kelley.

Given (Ambiguous): Starch devised a history text.

Improved: Starch published a history text.

He probably did devise some sort of a history text at some time in his life; the important thing to know is whether he "published" it.

18. Try to place modifiers (clauses, phrases, and words,) near the word or words they actually modify.

Given (Ambiguous): The young child reads words almost as readily inverted as upright.

Do the words "inverted" and "upright" refer to the "child" or to the "words" in the statement? The statement may mean the young child can read almost as readily resting on his head as standing on his feet. That may be true for some young children and false for others. The statement may mean the young child can read inverted words almost as readily as upright words. This should be either true or false in terms of reliable evidence.

Improved: The young child reads inverted words almost as readily as upright words.

19. The word *only* should be placed before or after the word it modifies. Misplacement of *only* may illogically express the implied contrast.

Given (Ambiguous): Injection of a hog only with serum confers an active immunity.

Which interpretation is correct?

*a.* A hog is the only animal upon which an injection of serum will confer an active immunity.

*b.* A hog and not some other animal. . . Implication being

that some active immunities require more than one specie of animal.

*c.* Injection of a hog with serum only, will confer an active immunity. Assuming the interpretation under (*c*), the statement would read:

Injection of a hog with serum only, confers an active immunity. Assuming the interpretation hnder (*c*), the

20. The correlatives *not only . . . . but also, either . . . or, neither . . . nor, both . . . and, on the one hand . . . on the other hand,* should be put before the words that they connect. Parallel construction seems to be preferred.

Given (Ambiguous): Morons neither are successful in their studies nor in their physical exercises.

Improved: Morons are successful in neither their studies nor physical exercises.

21. Do not omit a word, phrase, or clause necessary to the logical development of the thought.

False: An educational program should always precede a building program.

True: The consideration of an educational program should always precede that of a building program.

It was the latter thought which the instructor actually desired to express.

22. Avoid awkward sentence structures.

*a.* Interpositions of words.

Given: It is necessary, in order to insure the health of the children, to provide thirty cubic feet of fresh air per minute for each one.

Improved: In order to insure the health of children, it is necessary to provide thirty cubic feet of fresh air per minute for each child.

Still better: A child requires thirty cubic feet of fresh air per minute to insure health.

*b.* Dependence of clauses introduced by *who, which,* or *that.*

Given (very ambiguous and awkward): A poor reasoner takes unfruitful clues that lead to results that are unreliable, and that should warrant no serious thought.

The above statement is very poor and should not be used. Each clause is too "sweeping" in its content. Poor reasoners take some fruitful as well as probably many unfruitful clues that lead to all kinds of results, and so forth. It is better to eliminate such a statement rather than to attempt to improve it.

23. Avoid split prepositional phrases.

Given (awkward): The caloric theory of heat possessed the characteristics of, although experiments proved it incorrect, a scientific hypothesis.

Improved: The caloric theory of heat possessed the characteristics of a scientific hypothesis, although experiments proved it incorrect.

24. Avoid split infinitives.

Given (ambiguous and awkward): The natural sciences have tended to always advance as the qualitative distinctions have been reduced to quantitative.

Their advance has not been "always" in the sense of "kept on advancing." The advance is "as qualitative distinctions have been reduced to quantitative."

Improved: The natural sciences have always tended to advance as the qualitative distinctions have been reduced to quantitative distinctions.

25. Avoid weak and colorless words that destroy the emphasis in the statement. Rarely begin or end a sentence with general or indefinite words.

Given (ambiguous): All short cuts are only tricks of number and must be taught as such.

The word "such" is ambiguous. Does it mean "taught as such are taught," or does "such" refer to "short cuts" or to "only tricks of numbers"?

Improved: All short cuts must be taught as only tricks of number.

Given (weak beginning): There must be no unprepared dictation in the elementary school.

Improved: All dictation must be prepared in the elementary school.

25*a*. Try to develop the skillful use of climax in the true-false statement; i.e., the arrangement of words, phrases, or clauses, in an ascending series.

Given (good) Paleontology furnishes us with the only proof acceptable to scientists of the truth of the theory of organic evolution.

Improved: Paleontology furnishes scientists with the only acceptable proof of the theory of organic evolution.

26. Aim to develop skillful use of antithesis.

While the immediate occasion of our entrance into the World War was the sinking of the Lusitania and the Sussex, the underlying cause was our fear of German world domination.

27. Avoid the plural and the single reference to the same word in a statement.

Given (wrong and ambiguous): Statistical Method is one of the best, if not the best, method for research.

Improved: The statistical method is one of the best methods of research, if not the best.

In the first statement, the word "method" is used in both the plural and the singular senses The grammar is faulty.

An abbreviated form: The statistical method is the best method of research.

28. Avoid improper use of *when* and *where* clauses as predicate nominatives.

Given: Vitamin A is where an antirachitic substance is found in certain foods.

Improved: Vitamin A is the antirachitic substance found in certain foods.

Given: Clotting is when an insoluble substance in the last analysis is formed from a soluble substance.

Improved: Clotting is essentially a formation of an insoluble substance from a soluble one.

### *Grammar*

29. Avoid false syntax of any kind.

A sentence is ambiguously false in three senses. The student may legitimately declare the statement false in terms of (1)

grammar, (2) subject matter content, or (3) grammar and subject matter content. An examination in subject matter content is not intended to be an examination in grammar. Correct syntax limits a student to a decision of either true or false in the terms of the subject matter content only.

### *Diction*

Good DICTION is always essential. The ideal true-false statement should consist of words or digits whose definitions and usage are comprehended in one and the same way by everyone. It is probably true that such an ideal is seldom, if ever, realized. The unlettered and the highly educated use different sets of words to mean the same things. Word vulgarisms [3] are seldom associated with the selected diction of the formal written language. Between these two extremes is a group of words which do not purport to be elegant, formal, or precise. These blend into one extreme or the other with such continuity that ofttimes vulgarisms cannot be distinguished from colloquialisms, and colloquialisms cannot be distinguished from formal written language.

For example—Vulgarism: Let's stack 'em! Colloquialism: Let's eat! Formal language: Shall we go to lunch?

Again, words in the native language change in meaning and usage as time passes. The word "funny" referred to the "comical." However, "funny" is used to refer to things "strange" or "remarkable" as well as "comical." The word "quite" is used to mean "very," "somewhat," and "entirely."

He was quite well.
He was quite amazed.
This is quite enough.

### *Vocabulary Levels*

By "vocabulary levels" is meant changes of levels of words in terms of time.

Changes in a word both with respect to its level in vocabulary and in meaning and usage in terms of time, tend to destroy the possibility of everyone understanding a given word the same way.

[3] Royster, James Finch, and Thompson, Stith. *Manual and Notebook for English Composition.*

As shown in the chart, the word B in 1600 A.D. was used in the selected diction of formal language. In 1900 A.D. the same word was a vulgarism. Formal language would scorn its use. On the other hand, the word A advanced to be used as a colloquialism and the word C advanced into the realm of formal language usage.

CHART 8

CHANGES OF WORD USAGES BETWEEN FORMAL LANGUAGE AND VULGARISMS IN RELATION TO PERIODS OF TIME

| | | | | |
|---|---|---|---|---|
| Formal Language | B = fire,[4] darn[5] | C | C | C |
| Colloquialisms | C = singeth | B | | A |
| Vulgarisms | A[6] | A | B | B |
| | 1600 A.D. | 1000 A.D. | 1500 A.D. | 1900 A.D. |

When words are grouped to form a sentence, the possibility of the sentence being understood by all in the same way would seem to be subject to the degree of like understanding of the words, the phrases and the clauses which make up the sentence. The chances are that the degree of ambiguity of a single word will be less than the degree of ambiguity of combinations of words into phrases, clauses, and the like, which are known to vary in degree of comprehensibility. The degree to which one word may be ambiguous in meaning may be influenced by the degree to which another word is ambiguous in meaning. When the two words are used in a phrase, the resultant degree of ambiguity may be increased or decreased or even reduced to zero.

For example—consider the separate words: *one-half, statements, true, false, either or, the, of, and, are.* Analyze the following phrases and sentences for possible ambiguities:

(1) true-false
true and false
true or false
either true or false

(2) true-false statements
true and false statements
true or false statements
either true or false statements

(3) The statements are true-false
The statements are true and false

[4] Krapp, George Philip. *Modern English: Its Growth and Present Use,* pp. i-x, 358.
[5] Krapp, George Philip, *The English Language in America.*
[6] A is hypothetical.

The statements are true or false
The statements are either true or false

(4) One-half of the statements are either true or false.

(5) One-half of the statements are true and one-half of the statements are false.

It seems that the real and vital contributions of statement construction of the future lies in this unexplored and fruitful field of content analyses and relationships in statements concerning subject matter. Some techniques are available and only require application, whereas other techniques need to be discovered and developed.

Who knows but that a real and useful tool for curriculum research lies hidden in this challenging mass of subject matter.

It may be that the evolution of civilization has made the selections of things most worthwhile for the greater part in terms of those things most easily and homogeneously understood; that the things less homogeneously understood have been analyzed for content and made more and more homogeneously understandable when the need arose. Some of the content is so complex that it seems to defy analysis. Here the genius revels in speculation. Analysis of content of subject matter items might tend to validate subject matter items worthwhile in terms of basic behavioristic criteria.

30. Pay strict regard to the exact shade of meaning of each word used.

True: The content of an immoral play treats of immoral people.

The above statement is not criticized for word usage. The word "immoral" should be differentiated from "unmoral."

False: The content of an immoral play treats of unmoral people.

There probably is no person who is totally unmoral—without a single moral. *Unmoral* should not be used, when the intention of the instructor would probably necessitate the use of *immoral.*

31. Avoid slang. It would seem that interest in an examination could be stimulated by more dignified means.

Given: Parsnips are usually "jerked" in the fall in the northern part of the United States.

Improved: Parsnips are usually harvested in the fall in the northern part of the United States.

Given: When anabolism exceeds catabolism fatigue "gets goin'."

Improved: When anabolism exceeds catabolism, fatigue sets in.

32. Avoid the use of obsolete, archaic words or phrases. A dictionary should be consulted in all doubtful cases. Example: use *got,* not *gotten,* use *drunk,* not *drunken.*

33. Avoid the use of unusual or very difficult words, when a simpler form of expression may be used for the same thought.[7]

Given: Harmonization of color vibrations is extraneous to the mastery of hat construction.

Improved: Color harmony does not enter into the study of millinery.

34. Avoid improprieties—good English words used in the wrong sense, either as a wrong part of speech or as an incorrect meaning.

Given (Adjective used as a noun): All music students should take vocal for six months.

Improved: All music students should take vocal lessons for six months.

35. Avoid foreign words when a good English equivalent may be used (i.e., when the examination is in the English language). Certain foreign words are a part of the technical language of certain subject matter.

Logic—per se, a posterioni
Law—bona fide, nol pros, ex post facto
Politics—ultimatum, carte blanche
Fine Arts—rôle, technique, bas relief
Geometry—reductio ad absurdum
Chemistry—argentum, ferrum
Music—forte, pianissimo, crescendo

Given: Most tentative hypotheses may be reduced to *nada.*

*Nada* is a foreign word meaning *nothing.*

[7] This rule is not intended to include words or phrases in a statement specifically intended as subject matter words. For example, in physics: "difference in potential" and in general "potential progress." The word *potential* is used as a specific subject matter term in physics and in the second case as a general language term.

Improved: Most tentative hypotheses may prove to be worthless.

36. Do not use technical terms which do not belong to the subject matter upon which the examination is based.

Given: A mechanical equivalent of mental kinetics is unknown.

Improved: The average amount of achievement for a given amount of mental effort is unknown.

37. Express all statements in the third person. Use the first person only in case of direct quotation.

Given (Good): All contagious diseases are infectious.

Given: I said that enzymes are organic catalyzers.

Improved: Enzymes are organic catalyzers.

Given: In the metaphase of somatic mitosis, we have the reduction to one-half of the original number of chromosones.

Improved: The chromosones are reduced to one-half of their original number during the metaphase of somatic mitosis.

38. Do not use a word in more than one sense in the same statement.

Given (Ambiguous): In the true scientific thinking, biases are not scientific.

Improved: In true scientific thinking, biases play no part.

39. Avoid verbosity. Practice stating a thought adequately with a few necessary words.

Given (Ambiguous): Analyses of textbooks in arithmetic show that their authors are in close agreement on the content of arithmetic.

Improved: Analyses of the content of arithmetic texts show close agreement.

*Close agreement of content does* not necessarily mean *the authors* are in close agreement. The market for arithmetics may have partially dictated content in terms of what kind of an arithmetic textbook could be sold. It was the close agreement of content of arithmetic textbooks which was studied. It is very improbable that any kind of an analysis of the authors themselves would show them to be in agreement on anything.

Given: Opportunities for equalization of burden increase as the size of the taxing unit for support increases.

Improved: Opportunities for equalization of burden vary with the size of the taxing unit.

40. Avoid needless repetition of thought.

Given: Pupils are drilled too much on the fundamentals of arithmetic; they should be given less drill and more opportunity for reasoning.

Improved: In arithmetic, pupils should be given less drill on fundamentals and more opportunity for reasoning.

The student's decision concerning "pupils are drilled too much on fundamentals of arithmetic" determines his decision concerning "they should be given less drill." If the first part of the statement were *false* then the phrase "more opportunity for reasoning" plays no part in the decision for the statement as a whole. If the first part of the sentence were *true* then the phrase "more opportunity for reasoning" plays a part in the decision for the statement as a whole. In either the *true* or the *false* case, the phrase "they should be given less drill" plays no significant part in the decision for the statement as a whole.

41. Avoid redudancy—non-essential words, phrases, or clauses.

Examples of total or partial repetition of word meaning:

*figure,* not *figure up*
*open,* not *open up*
*over,* not *over again*
*off,* not *off of*
*whence,* not *from whence*

Examples of repetition of construction in grammar:

Double negatives.

Given: Dictation should not be unprepared in the elementary school.

Improved: Dictation in the elementary school should be prepared.

Examples of awkward use of *because*

Given (Ambiguous): The reason lyric poetry is called lyric poetry is because it appeals to the emotions.

Improved: Lyric poetry is so called because it appeals to the emotions.

This type of poetry probably is intended to make an appeal to the emotions. It is doubtful if the entire range of individual emotional differences in humans would be partly or wholly affected by either the best or the worst poetry which may be called lyric. It is doubtful if any given lyric poem would appeal to the emotions of all people. The statement in its better form is probably false.

Restatement in a qualified form. According to ——, good lyric poetry appeals to the emotions (—— would be an author, the opinion of some person or group, or scientific evidence.)

42. Avoid inexact qualifying phrases such as *sort of, as it were, pretty near, nearly, kind of, not quite, so so,* and so on.

Given (Ambiguous): Mercuric chloride is a sort of disinfectant.

Improved: Mercuric chloride is an efficient disinfectant.

43. Avoid an overuse of superlatives.

Given: Periods of quickest or most rapid growth are periods of greatest dullness in pupils.

This statement would mean the same without the word *quickest.*

44. State a quantitative fact quantitatively.

Given (Ambiguous): Geometry marks vary less than English or history marks.

How much less? If the *less* is not significant, then the chances are that the statement is doubtful as to whether true or false. If the *less* is significant, then the amount of the significance is known by statistical methods. For instance, the Critical Ratio [8] formula or the Experimental Index [9] may be used to determine the degree of significant differences between two arrays of independent variables.

An improved statement is: The value of the experimental coefficient determined from distributions of per cents in geometry and English or history respectively, was three in favor of geometry.

45. Avoid unpleasant effects, such as an excessive succession of like sounds. (This sentence itself illustrates the fault.)

[8] McGaughy, J. R. *The Fiscal Administration of City School Systems,* p. 55.
[9] McCall, Wm. A. *How to Experiment in Education.*

Given: Transportation is a foundation for unification of a nation.

Improved: Transportation is a fundamental factor in national unity.

46. Try to select the right word and put it in its place. Certain feelings become attached to certain words through usage. Some words are humorous; others are ugly; some are very formal; while still others are poetical.

Given (Good): All bacteria are eliminated at 0° Centigrade.

Improved: All bacteria are destroyed at 0° Centigrade.

Given: Energy is the ability to perform work.

Improved: Energy is the capacity to do work.

### VALIDATION BY EXPERT JUDGES

*Procedure*—In order to justify the procedure followed in this chapter, 125 true-false statements were submitted to a group of judges in a given subject matter. At least 4, and not more than 5, members composed the group of judges revising the statements.

Session No. 1: Five judges judged true-false statements for 1 hour. Three of the judges were experts in the subject matter and two were advanced graduate students. One subject matter expert had an appointment and left at the end of the hour. Four judges remained through the 2½-hour session. An expert stenographer recorded the discussion of the entire session.

Session No. 2: At a second sitting, four judges revised statements for three hours. Two of these were experts in the subject matter. One of these judges did not attend session No. 1. The remaining two judges were the advanced graduate students who attended session No. 1. No stenographic record was taken during the three-hour session.

One of the two graduate student judges prepared the 131 true-false statements. This student was conscientious and did his best to submit statements which were either true or false. He answered each of the 125 statements as either true or false before submitting them to the judges. This was evidence of what he thought was the answer to each statement.

The problem in each session was: What is the basis of each decision for each true-false statement?

CHART 9

AN INVENTORY OF REASONS GIVEN BY EXPERTS IN THE SUBJECT-MATTER FOR ACCEPTANCE OF THE 55 TRUE-FALSE STATEMENTS WITHOUT REVISION, WITH AN EXAMPLE FOR EACH REASON

| Reasons for Acceptance of the Statement Without Revision | *Examples of statements accepted without revision* |
|---|---|
| Clear Statement | Opportunities for equalization of burden vary with size of the taxing unit. |
| Brief Statement | Public Education per average daily attendent tends to the cost less in a rich community. |
| True in Terms of the Best Known Evidence | Boards of education of city school systems should be fiscally independent of Municipal control. |
| False in Terms of the best Known Evidence | The best system of heating and ventilation is the direct mechanical. |
| Concise Statement | Supervision of an expert teacher is unnecessary. |
| Accurate Statement | Approximately 68.26 per cent of the cases in a normal distribution are within one sigma of the average. |
| Definite Statement | The possible range of the Pearson product-moment coefficient of correlation is from 0 to 1. |

TABLE 24

TENDENCY OF DISTRIBUTIONS OF TRUE-FALSE STATEMENTS TOWARD A CORRECT ANSWER OF FALSE IN THREE LAW COURSES WHEN ALL STATEMENTS CONTAINING THE WORD *NOT* ARE STATED WITHOUT THE WORD *NOT*

| Course in Law | Correct Answer | Statements Containing | | | All Statements Containing *always* and *never* Words Stated as Being *always* True |
|---|---|---|---|---|---|
| | | *Always* and *Never* | The word *not*, but the word *not* omitted | *Always* and *never* Words Other Than the Word *not* | |
| Trusts | True | 17 | 13 | 1 | 14 |
| | False | 15 | 15 | 3 | 18 |
| Torts | True | 29 | 23 | 2 | 25 |
| | False | 29 | 27 | 6 | 33 |
| Business Law | True | 18 | 13 | 1 | 14 |
| | False | 19 | 17 | 6 | 23 |

CHART 10

INVENTORY OF THE REASONS GIVEN BY EXPERTS IN THE SUBJECT MATTER FOR REVISION OF FIFTY-ONE TRUE-FALSE STATEMENTS, WITH A STATEMENT, ITS REVISION, AND THE NUMBER OF THE VALIDATING RULE

| Reason Given for the Revision of the Statement | Statement Before Revision | Revised Statement | Number of the Validated Rule |
|---|---|---|---|
| Misspelled Word ..... | When a community is divided on a *controuersial* issue, the *uperintendent* should in all cases *efinitely* take one side or the other. | When a community is divided on a *controversial* issue the *superintendent* should in all cases *definitely* take one side or the other. | 1 |
| Loose Sentence Structure | McLure found that humidity was more important than temperature and that circulation of air was more important than either in preserving the health of pupils. | 1. McLure found that humidity was more important than temperature in preserving health.<br>2. McLure found that circulation of air was more important than either humidity or temperature in preserving the health. | 7 |
| Dangling Clause ..... | Members of the State Board of Education should be paid a salary which is a living wage, if elected by the people. | Members of the State Board of Education should be paid a salary which is a living wage. | 15 |
| Indefinite Pronoun *It* ... | It is dangerous for the state legislature to determine the content of curriculum. | The state legislature should have the power to determine the content of curriculum. | 12 |
| Clarity of Thought ..... | Approximately 68 per cent of the cases in any distribution fall between negative sigma and positive sigma. | Approximately 68 per cent of the cases in a normal distribution are within one standard deviation of the average. | 8 |
| Omission of Necessary Phrase ..... | An educational program should always precede a building program. | The consideration of an educational program should always precede a building program. | 21 |
| Antithesis-Contrast ..... | When a charge should be technically divided between maintenance and capital outlay, it is customary to charge it all to maintenance. | Although sound theory may require the division of a certain expenditure, in practice the whole expenditure is charged to maintenance. | 26 |
| Awkward Sentence ..... | Very little money should be spent for publicity except for unusual situations such as bond campaigns. | 1. It is ethical for a Board of Education to expend public money for school publicity.<br>2. School publicity should be continuous rather than intermittent. | 22 |

| | | | |
|---|---|---|---|
| Placement of Modifiers .. | Primarily a State Board of Education should be a legislative body. | A State Board of Education should be primarily a legislative body. | 18 |
| Disagreement in Number | The differences among states in ability to provide educational opportunity is negligible. | The differences among states in ability to provide educational opportunity are negligible. | 29 |
| Unusual Word ..... | The greater part of the work of the Department of Education, as provided in the new Bill as proposed at the Indianapolis N. E. A. July, 1925, is *ostensibly* in the field of educational research. | The greater part of the work of the Department of Education, as provided in the New Bill, as proposed at the Indianapolis N. E. A. July, 1925, would be in the field of educational research. | 33 |
| Word Used in Wrong Sense | It is more economical for large cities than small cities to carry their own insurance risks on buildings and equipment, than to insure in private companies. | Substitute the word "defensible" for the word "economical." | 34 or 46 |
| Verbosity ..... | The assessments levied against teachers in the interest of providing for retirement salaries should be kept in separate accounts to the credit of the individual teacher, similar to a savings bank account. | All contributions to state retirement funds should be kept in separate accounts to the credit of the individual teacher. | 39 |
| Redundancy of Phrase .. | State adoption of all textbooks is considered advisable by the majority of educational experts. | State adoption of all textbooks is advisable. | 41 |
| Weak Beginning of a Statement ..... | There is a conflict between the awarding for effort and the awarding for ability in state aid to local school systems. | State aid to local school systems involves a conflict between the payment for effort and the payment for need. | 25 |
| **Redundancy—Double Negative** ..... | It is not financially unsound to bond for current expenses. | It is financially sound to bond for current expenses. | 41 |

CHART 10—(*Continued*)

REASONS GIVEN BY EXPERTS FOR ELIMINATION OF TRUE-FALSE STATEMENTS IN THE GIVEN SUBJECT MATTER

| Reason Given for the Revision of the Statement | Statement Before Revision | Revised Statement | Number of the Validated Rule |
|---|---|---|---|
| Poor Suggestion of a Word | Cubberley believes that a State Board of Education should be composed of five members, one elected each year by the state at large. | Cubberley believes that a State Board of Education should consist of five members, one elected each year by the state at large. | 46 |
| Statement Poorly "Limited" in Scope ..... | The Federal land grants were largely responsible for the establishment of a public school system. | According to Swift the Federal land grants gave the impetus to the establishment of a public school system west of the Alleghany Mountains. | 21 |
| One Idea in Each Statement ..... | All members of the staff should be directly responsible to the superintendent of schools. He, in turn, is directly responsible to the Board of Education, which is directly responsible to the people. | All members of the staff should be directly responsible to the superintendent. | 9 |

CHART 11

REASONS GIVEN BY EXPERTS FOR ELIMINATION OF TRUE-FALSE STATEMENTS IN THE GIVEN SUBJECT MATTER

| Reason Given for Elimination | The Eliminated Statement |
|---|---|
| Insignificant Statement ..... | Twenty-seven national organizations advocate the establishment of a federal Department of Education. |
| Self-evident Statement—too easy ..... | 1. A state board of education should be primarily a legislative body.<br>2. All states should have a social and economic interest in the provisions for education in every other state. |
| Answer to One Statement Suggested or Contained in Another Statement ..... | 1. The county superintendent of schools should be appointed by the county board of education and serve at their pleasure.<br>2. (Omitted) The county superintendent should be chosen by the people. |
| Too Many Statements Devoted to Particular Part of Subject-Matter of the Course... | (Three statements upon the topic of school publicity were included in the examination. It was suggested that was a sufficient and representative number). |
| Very Controversial Statement—Lack of Evidence, Conflict of Authority, Conflict of Philosophy or Theory ..... | 1. There should be constitutional limitations upon the rate of taxation for education in every state.<br>2. Transportation is largely responsible for increased attendance of continuation schools over the old district schools.<br>3. If the state inspects non-tuition private and parochial schools and sets up minimum requirements for them, it should give aid to them.<br>4. Indefinite tenure is the most satisfactory form of tenure.<br>5. A school superintendent should have two years of professional training.<br>6. The most important measure of a school system may be found through the following question: "How much have the pupils learned?"<br>7. The principal should give the largest proportion of his time in supervision of the poorest teachers.<br>8. The most of the work of the committee on any one subject of the curriculum has actually been the opinion of one man rather than group judgment or survey of judgment.<br>9. The people who have children in school and who pay for the public school system should determine what should be taught in it. |
| Ambiguous Statement in Need of Restatement ..... | 1. Pie graphs are measures on the circumference of a circle.<br>2. Area is commonly violated when increase or decrease is represented by cartoons. |

No statement was finally accepted unless all the judges agreed that the statement was either *true* or *false*.

*Results.*—During session No. 1 a unanimous decision was rendered on 40 statements in 2½ hours.

During session No. 2 a unanimous decision was rendered on 85 statements in three hours.

The increase in rate of judging the statements was thought to be due to the improvement in the judges' technique of judging statements.

A summary of the results of this method when applied to the 131 statements was:

55 statements, not revised in sentence structure.
45 statements, partially revised in sentence structure.
6 statements, restated.
2 statements, added to the examination.
25 statements, eliminated from the 131.

The reason for the disposition of each statement was stated by the subject matter expert.

*Unrevised statements.*—Chart 9 below shows an inventory of the reasons stated by expert judges for accepting each of 55 true-false statements without revision. These reasons overlap one another more or less. It would seem that an acceptable true-false statement is a sentence which is clearly, concisely and accurately expressed in terms of the best known evidence.

*Revised statements.*—Chart 10 shows the reasons given by expert judges in the subject matter for revisions of true-false statements. Their reasons are not unlike those given by the empirical rules.

There is a very close relationship between the actual process of revising true-false statements by group judgment and the rules of good English composition.

*Additional statements.*—The reason given for including the two new statements was:

Every phase of a subject matter which is judged to be of significant importance should be proportionately represented in a true-false examination.

*Eliminated statements.*—Chart 11 shows a brief inventory of the bases upon which a given true-false statement may be eliminated from an examination. The important revelation

of this chart is the fact that controversial subject matter content is eliminated from true-false examinations.

This controversial subject matter is seldom judged to be either true or false in terms of the amount and best evidence available. This would seem to be a limitation of the true-false examination.

Chapter X discusses this problem and proposes a tentative solution, which will require further research.

## CONCLUSIONS

The results of the expert judgment technique and the empirical technique described at the opening of this chapter agree. The empirical technique may be considered as partially validated by means of the expert judgment technique. The latter technique is an actual process of constructing true-false statements.

The evidence indicates that grouping of expert judgments is an improvement upon the labors of one competent and careful graduate student in the subject matter. In the absence of conference method a set of rules would probably be the next best device.

The grouped judgment method uncovered a group of supposedly true-false statements which were eliminated. These statements, in part, apparently represent a body of knowledge which is not readily formulated into true-false statements. This knowledge is admittedly controversial. Either evidence is lacking, or authorities conflict in their published or oral utterances.

The evidence indicates that the true-false statement probably cares for factual knowledge in so far as that knowledge may be judged to be either true or false. It is further evident that the true-false statement is somewhat limited in its nature with respect to that body of subject matter which is on the border line of being known or unknown. When the subject matter of a given statement is doubtful as to whether true or false, that statement ceases to be a true-false statement in the sense that it may be judged by competent judges to be consistently either true or false. Chapter X attacks the disposition of this problem.

# CHAPTER VIII

## SPECIFIC DETERMINERS IN TRUE-FALSE STATEMENT

### INTRODUCTION

*Definition.*—A *specific determiner* may be defined as a word (may also be a group of words) in a true-false statement, which tends to cause that statement to be more times *true* than *false* or more times *false* than *true*.

In the abstract, the random guessing of statements containing a given specific determiner will give a score value other than zero, i.e., $R - W = \pm K$.

Examples of a specific determiner are:

1. The statements which contain the idea that the fact is *always* or *never* true.

*All* sheep have wool.

The word *all* is the specific determiner.

2. The statements which contain the idea of degree or comparison.

The elephant is *larger than* the mouse.

The word phrase *larger than* is the specific determiner.

*Scope of the problem.*—It is possible that the number of different specific determiners in the English language is limited only by the number of words in the language plus a certain large number of combinations and permutations of those words. This field alone seems limitless to the writer.

### STATEMENT OF PURPOSES

The purposes of this chapter are:

1. To discover certain specific determiners groups as they occurred in true-false examinations.
2. To demonstrate their effect upon the total score of the examination.
3. To suggest how any specific determiner may be controlled.

SPECIFIC DETERMINER GROUPS

*Procedure* (*words and phrases*).—Four general groups of words and phrases were studied.

Group I: *always* and *never* (also the word *not*). (See Appendix G, for the inventory used in this study.)

Group II: *Almost always, almost never, in general,* etc. (See Appendix G, for the inventory used in this study.)

Group III: *Equal to, congruent to, similar to, equivalent to* and *same size as.*

Group IV: *Degree or comparison—more than, less than, -er than, the more . . . the more, the less . . . the less, the more . . . the less,* and *the less . . . the more.*

*Results*

1. Among 1950 true-false statements, there were 325 *always* and *never* statements. Of these 212 were *false* and 113 were *true.* The statements were collected from courses in Contemporary Civilization, Government, Economics, Plane Geometry, and Physics, in Columbia College, New York City.

2. Table 24 indicates the tendency of the word *not* in true-false statements in three Law courses toward *false* as the correct answer.

3. Among 218 *almost always* and *almost never* statements 106 were *false* and 112 were *true.* The course was Contemporary Civilization, made up of eight to ten other courses in Columbia College.

4. Among 212 *equal to, congruent to* statements in Plane Geometry, 106 were *true* and 106 were *false.*

5. Among 1776 true-false statements, there were 130 *degree or comparison* statements. Of these 83 were *true* and 47 were *false.* The courses used were Contemporary Civilization, Economics, Government, and Plane Geometry given in Columbia College.[1]

*Conclusions.*—In the above examinations:

1. *Always* and *never* statements tend to be *false.*
2. *Degree or comparison* statements tend to be *true.*

*Procedure* (*Dependent Clauses*).—An analysis of the type of dependent clauses in complex true-false statements was made in Contemporary Civilization, Economics, and Government.

[1] Complete data are given on pages 128-33 of the manuscript copy of this dissertation on file at Teachers College Library.

*Results.*—Table 28 reveals two new facts:

1. Very close to 26 per cent of all dependent clauses are either *noun* or *adjective* clauses introduced by *who, that,* or *which.*

TABLE 28

DISTRIBUTION OF 604 DEPENDENT CLAUSES IN COMPLEX TRUE-FALSE STATEMENTS FROM CONTEMPORARY CIVILIZATION, ECONOMICS, AND GOVERNMENT, BASED ON 1509 STATEMENTS

| Type of Dependent Clause | Correct Answer | Contemporary Civilization | Economics | Government | Total |
|---|---|---|---|---|---|
| Noun and Adjective | True | 41 | 30 | 31 | 102 |
| | False | 51 | 48 | 31 | 130 |
| Concession | True | 16 | 4 | 9 | 29 |
| | False | 14 | 1 | 4 | 19 |
| Result | True | 4 | 0 | 0 | 4 |
| | False | 9 | 1 | 1 | 11 |
| Purpose | True | 6 | 0 | 2 | 8 |
| | False | 4 | 1 | 1 | 6 |
| Condition | True | 4 | 15 | 2 | 21 |
| | False | 6 | 10 | 3 | 19 |
| Cause or Reason | True | 25 | 7 | 6 | 38 |
| | False | 48 | 25 | 12 | 85 |
| Manner | True | 1 | 1 | 1 | 3 |
| | False | 1 | 0 | 0 | 1 |
| Degree or Comparison | True | 29 | 8 | 2 | 39 |
| | False | 12 | 9 | 0 | 21 |
| Place | True | 1 | 2 | 2 | 5 |
| | False | 0 | 1 | 3 | 4 |
| Time | True | 17 | 10 | 7 | 34 |
| | False | 10 | 12 | 3 | 25 |
| Total ..... | | 299 | 185 | 120 | 604 |
| Total Statement Basis ... | | 683 | 533 | 293 | 1509 |

2. Among 123 *cause* or *reason* clauses 85 were *false* and 38 were *true.*

*Conclusion.*—*Cause* or *reason* clauses tend to be *false* in the above examinations.

*Independent clauses.*—Only 79 compound statements occurred in a total of 1,509 true-false statements. These tended to be as often *true* as *false*.[2]

*Results (simple statements).*—In Table 30 a brief study of the simple sentence was made for Economics and Government. *True* statements seem to be construed a little more often than *false* statements.

*Conclusion.*—Probably *true* statements tend to be more often constructed than *false* statements.

TABLE 30

TRUE-FALSE STATEMENT DISTRIBUTION OF 437 SIMPLE STATEMENTS IN ECONOMICS AND GOVERNMENT

| Correct Answer | Simple Statements in Economics | Government | Total |
|---|---|---|---|
| True | 156 | 76 | 232 |
| False | 137 | 68 | 205 |
| Total | 293 | 144 | 437 |

RELATION OF SPECIFIC DETERMINERS TO SCORE

In brief these facts would tend to say that in a non-mathematical subject matter

(*a*) 2 out of 3 *always* and *never* statements are *false*.

(*b*) 2 out of 3 *degree or comparison* statements are *true*.

(*c*) 2 out of 3 *cause* or *reason* clauses are *false*.

Let us consider that the student knew nothing about the subject matter and used only his knowledge of *specific determiner* groups.

On the two out of three basis his score on an entire examination of 151 statements could be computed as follows:

Assume

21 *Always* and *Never*
15 *Degree* or *Comparison*
21 *Cause* or *Reason Clauses*
57

151 total number of statements in the examination
57
94 statements of which the student knew nothing.

[2] Table XXIX is given on page 136 of the manuscript copy of this dissertation, Teachers College Library.

On the 57 statements his score would be

$57 \times \frac{2}{3}$ right $= 38$ right
$19$ wrong
Score $= 19$

On the 94 statements his score would be

$94 \times \frac{1}{2} = 47$ right
$47$ wrong
Score $= 0$

Total score for the examination $= 0 + 19 = 19$.

Theoretically when an individual knows nothing about a subject matter, the score is zero. If *specific determiners* like the above are present and the student knows about their tendencies toward either *true* or *false,* a zero score would be improbable even though he knows nothing of the subject matter.

### THE CONTROL OF SPECIFIC DETERMINERS

1. Formulate about as many statements *true* as *false* which contain any one group of *specific determiners.* It does not seem advisable to recommend exactly one-half. This would aid the student if he knew the answer to even a few of the statements contained in a given specific determiner group.

For example: Consider the *always* and *never* and *degree* or *comparison* groups of specific determiners:

construct 11 *always* and *never* as *true*
and 4 *always* and *never* as *false*
then construct 14 *degree* or *comparison* as *false*
and 11 *degree* or *comparison* as *true*

This keeps away from their tendencies to be predominantly either *true* or *false.*

2. Avoid the use of the word *not* in order to change a statement from *true* to *false* or from *false* to *true.*

Example:
Statement—Most fish do not lay eggs.
True—Most fish do lay eggs.
Therefore False—Most fish do not lay eggs.

3. Do not avoid *always* and *never* statements. Brinkley [3] says, "In general, it is wise to avoid questions containing the

[3] Brinkley, Sterling G. *Values of New Type Examinations in the High School,* p. 38.

statement that a given fact is *always true* or *never happens.*" In the last analysis, his recommendation is undoubtedly based upon a conviction that such statements are usually, or in general false.

This determiner group may be controlled.

Formulate *always* and *never* statements so that about one-half are *true* and about one-half are *false.*

Avoid using more than one specific determiner of falseness in one statement.

Example:

Poor—Chapman and Count regard the "complex" as distinctly pathological and insist that it has no place in the life of the normal individual.

The first clause is *false,* due to the phrase "distinctly pathological."

The second clause is also *false,* due to the word "no."

Two determiners are in the statement and both are contrary to the facts as published.

A student need know only one, and either one, to answer the above statement as *false.*

5. Avoid any statement the answer to which is contained in or suggested by another statement in the same examination.[4]

Examples:

83. The transfer effect of school subjects is much higher than is usually supposed.

100. The fact of transfer is open to serious doubt.[5]

## CONCLUSION

Specific determiners may be controlled if the number of *true* and the number of *false* statements in a given specific determiner group are approximately equal to each other.

[4] *Op. cit.,* p. 161.

[5] Examples and details and described on pages 145-49 of the manuscript copy of this dissertation, Teachers College Library.

# CHAPTER IX

## WORD-LENGTH OF THE STATEMENT

### PRESENT PRACTICE

A true-false statement may (1) be "good" or "not good"; (2) have any number of words; (3) have any kind of sentence; (4) have any degree of reading difficulty independent of the subject matter being tested.

### STATEMENT OF THE PURPOSES

The purposes of this chapter are:

1. To define "goodness" of a statement.
2. To relate "goodness" to word-length of a statement.
3. To simplify true-false statements intended to test either "verbal memory" or general "reasoning ability" in a subject matter.

### DEFINITION OF "GOODNESS"

What is meant by a "good statement? The best reference by Wood [1] is quoted as follows:

*Validity of a Statement.*—The preceding remarks are made partly to call attention again to the great importance of a constructive attitude toward the essay type of examination, and partly to introduce the data which have been collected on the "goodness" of each of the 872 statements in five true-false law examinations. Each of these questions was tested in the following manner: All the students in, say, Course 1, were arranged in the order of their final grades for Course 1. The final grades were assigned by averaging the grades on the Essay and New test parts of the final examination. The lowest 15 per cent and the highest 15 per cent of the students were thus separated into the "Poor" and "Good" groups. It was then ascertained what percentages of these two groups of students marked correctly each question in the Course 1 New test. If, let us say, 20 per cent of the Poor group marked question 1 correctly and 90 per cent of the Good group marked it correctly, then it is considered a good question.

[1] Wood, Ben. D. "Measurement of Law School Work, II," *Columbia Law Review.*

But if 80 per cent of the Poor and only 20 per cent of the Good group marked it correctly, then it is a bad question, since, in so far as it has any influence at all, it tends to displace the scores of good students downward and of poor students upwards;—it tends to mix the goats with the sheep, whereas the purpose of an examination is to separate them. Having found the percentages of the Poor and Good groups, respectively, marking each question correctly, the former was subtracted from the latter. If a question is a good one, the difference will be positive; but if the number of poor students getting it right is greater than the number of good students getting it right, then the difference is negative. How good or bad a question is may be roughly judged by the magnitude of the difference, positive or negative, respectively. Any negative difference, or any zero or near-zero difference, indicates that a question is either bad or "dead," i.e., it either tends to mix the goats with the sheep, or has no systematic influence at all in classifying the students.

### CORRELATION OF "GOODNESS" AND WORD-LENGTH

*Definition.*—Degree of "goodness" of a statement is the per cent of correct response by the best one-fifth of the group of students minus the per cent of correct response by the poorest one-fifth of the group of students. (Degree of "goodness" may also be expressed as a ratio, the denominator of which is a constant.[2])

*Procedure.*—1. Statements were selected which had a 40 per cent to 80 per cent correct response for the middle one-fifth of the entire group of students.

2. The twenty-seven available Law[3] statements whose degree of "goodness" was more than + 0.25 were correlated with their respective word-lengths. (Call these the "good" statements.)

3. The thirty-four available Law statements whose degree of "goodness" was less than + 0.25 were correlated with their respective word-lengths. (Call these the "poor" statements.)

4. The word-lengths of the 27 statements were compared with the word-lengths of the 34 statements for differences in the word-lengths of their averages.

5. Tables 32 and 34 display the facts of this study.

*Results.*—1. The Pearson correlation coefficient[4] of the data in Table 32 is ±0.13. The ±4 P. E. range of the $r$ is +0.70

[2] Pages 151-52 of the manuscript copy of this dissertation, Teachers College Library, give full discussion of this.
[3] Final examination in Torts, 1924, Columbia University.
[4] Yule, G. U. *An Introduction to the Theory of Statistics.*

to —0.35. The Predictive Index[5] of +0.70 is 0.29, which is a little better than chance.

2. The Pearson coefficient of the data in Table 34 is ±0.05. The ±4 P. E. range of the $r$ is +0.49 to —0.39.

3. The Critical Ratio between the mean word-length of the 27 statements (Table 32) and the mean word-length of the 34 statements (Table 34) is 0.94.

4. Sixty-eight per cent of the statements in Table 32 have a word-length range between 11 and 25 words.

5. Eight of the ten best statements have degree of "goodness" values ranging between +0.62 and +0.81. The word-length range of these eight statements is from 13 to 20 words, inclusive. The Median word-length is 16 words.

*Conclusions.*—The evidence indicates that:

1. Little relationship exists between "goodness" and word-length of a statement.

2. The average word-length of "good" and the average word-length of "poor" statements are not significantly different.

3. A "good" statement can be constructed whose word-length is between 10 and 20 words. "Good" statements, however, are acceptable independent of word-length.[6]

### THE TRUE-FALSE STATEMENT IN RELATION TO "VERBAL MEMORY" OR VERY LITTLE GENERAL "REASONING ABILITY"[7]

*General "reasoning ability."*—It is impossible for the writer to explain what is meant by a fact which is so isolated that some small part of general "reasoning ability" is not used. What general "reasoning ability" is, likewise defies analysis. The next best approach is by concrete example.

A fact probably involving relatively little general "reasoning ability" is:

> New York City is located in New York State.

A fact probably involving a relatively larger amount of general "reasoning ability" is:

> An adequate education would free us from the trial and error process of learning.

[5] Bailor, E. M. *Content and Form in Tests of Intelligence*, pp. 27-28.
[6] Complete date are given on pages 150-59 of the manuscript copy of this dissertation, Teachers College Library.
[7] Wood, Ben. D. *Op. cit.*, p. 7.

*Procedure.*—An examination of 160 items in the subject matter of Anatomy was selected. An analysis of word-length and sentence structure was made. It was designed to test specific (verbal memory) information.

*Results.*—1. The average word-length for 154 statements was 14 words.

2. Simple declarative sentence was used in 151 of the 160 statements.[8]

### THE TRUE-FALSE STATEMENT IN RELATION TO GENERAL "REASONING ABILITY"

Wood[9] reports on some examinations which "were designed to measure not merely legal information and rules of law, but also 'reasoning ability' and 'power to apply knowledge of the law successfully to complicated and knotty legal problems.' That these new type law examinations did measure this 'reasoning ability' better than old type law examination is amply shown in the analyses of the results."[10]

*Procedure.*—1. An examination of 160 true-false statements in the Law course of Pleading and Practice was selected. An analysis of word-length and sentence structure was made.

*Results.*—1. The average word-length for 75 complex declarative statements was 30 words.

2. The complex declarative sentence was used in 130 of the 160 statements.[11]

### ANALYSIS OF PROBLEM TYPE OF TRUE-FALSE STATEMENT, USING COMPLEX AND SIMPLE SENTENCES

*Procedure.*—Another examination of 161 statements in Torts was selected by Wood as devised to test general "reasoning ability." This examination involved two distinct types of statement.

(See Appendix H, for examples of the problem type of true-false statement.)

[8] Complete data are given on pages 159-61 of the manuscript copy of this dissertation, Teachers College Library.
[9] Wood, Ben. D. "Studies in Achievement Tests." *Journal of Educational Psychology*, January, 1926.
[10] Wood, Ben D. "Measurement of Law School Work." *Columbia Law Review*, March, 1924, and March, 1925.
[11] Complete data are given on pages 162-64 of the manuscript copy of this dissertation on file at Teachers College Library.

*Results.*—1. The average word-length of 58 complex declarative sentences was 25 words.

2. The average word-length of the 85 simple declarative sentences of the problem true-false type was 14 words.[12]

## SUMMARY OF CONCLUSIONS

1. In present practice little relationship exists between "goodness" and word-length of a statement.

2. The simple declarative sentence should be used to test specific (verbal memory) information.

3. Wood [13] reports a Pearson correlation coefficient of validity for the Torts examination of this study of +0.82.

Therefore the problem true-false type of statement is a means of simplifying and shortening sentence structure. General "reasoning ability" may be tested by means of the simple declarative sentence introduced by a problem. (See Appendix H, for sample.)

4. The simple declarative sentence has an average word-length of 14 words.

5. "Good" true-false statements can be constructed whose word-lengths are 10 to 20 words.

[12] Complete data in manuscript copy, pages 164-70.
[13] Wood, Ben D. "Studies in Achievement Tests." *Journal of Educational Psychology*, Vol. XVII, No. 1, p. 12, January, 1926.

# CHAPTER X

## NATURE OF THE TRUE-FALSE STATEMENT

### ASSUMPTIONS

It is assumed that the responses to a true-false statement are limited to *true* (+), *false* (0), and the *omission* (no mark).

It is further assumed that the statement is either true or false in terms of the criterion of acceptability (See Chapter XI) and not doubtful as to whether the statement is *true* or *false* in terms of the same criterion of acceptability.

*Criterion of acceptability.*—The criterion of acceptability in the true-false examination as it is used to-day, in the last analysis, is the acceptance of the statement as either *true* or *false* by the instructor or staff of instructors of the given subject matter. Some better basis needs to be developed, in order to relieve the student of responding to a statement against his convictions. The idea which Kilpatrick uses in his true-false examinations for his courses in Philosophy of Education seems to offer help in this direction.

His instructions to a student are to this effect: "If you do not believe a statement is *true* but you feel that the instructor thinks the statement is *true,* mark the statement *true* and place the letters, 'fpx,' to the lest of the statement ('fpx' means 'for purposes of the course')."

However, the instructor or staff of instructors should formulate their decision into a statement based on opinion, on the factual evidence which is available, on the logic of the situation as a consequence of certain assumed hypotheses, or on any other basis which may be valid for the purpose of a decision.

### THE TRUE-FALSE STATEMENT AND INFORMATION (MEMORY OF A FACT)

The evidence presented in Chapter IX indicates that a true-false statement may be either *true* or *false* on a simple subject

matter item (verbal memory) information basis. It may be used in any subject matter in which there is a verbal memory content. Verbal memory content in subject matter may be illustrated by the following concepts: dates; names of persons, animals, plants, minerals, elements, compounds, cities, lakes, rivers, mountains, etc.; events; titles of persons, books, or publications; direct quotation of words, phrases, clauses, sentences, verses, paragraphs; weight; money values; etc. (See Chart 13.)

### THE TRUE-FALSE STATEMENT AND GENERAL "REASONING ABILITY"

The evidence in Chapter IX also indicates the nature of a true-false statement which is intended to test the general "reasoning ability" of the student in a given subject matter. So far as language will permit, the statement must not be ambiguous. Statements which may be interpreted in more than one way often may legitimately be answered as *true* by one interpretation and *false* by the second interpretation. (See Chapter XI.)

In order for the reasoning to be direct and clear-cut on the part of the student the evidence seems to indicate that.

1. The *words* of the statement should be as direct, concise, simple, and accurate as possible. This is assumed, for reading comprehension, per se, probably should be tested as little as possible when subject matter other than reading comprehension is being tested.

2. The statement may contain a certain amount of evidence recognized by the student as a basis for further reasoning.

3. The subject matter evidence not contained in the wording of the statement, but necessary to answer the statement must be accurately and clearly associated with the evidence printed in the statement. If the only evidence available which the student may recall and associated with the printed evidence in the statement is, itself, *indeterminate* (that is, doubtful as to whether *true* or *false*), then the response to the statement is indeterminate. Such a statement is not a true-false statement in the sense assumed at the beginning of this chapter.

It seems probable that certain aspects of "reasoning ability" [1] can be tested by a true-false statement when the evidence printed in the statement and the essential associated evidence not printed

[1] Wood, Ben D. "Measurement of Law School Work," *Columbia Law Review*, Vol. XXV, No. 3, March, 1925.

in the statement are not susceptible to interpretations which would make the statement doubtful as to whether *true* or *false*. A statement which is not susceptible to more than one acceptable interpretation may be defined as an objective statement.

The above does not mean that a correct answer to a statement is possible in terms of only one specified series of related subject matter items. The correct answer may be arrived at by any one of several groups of related subject matter items. It does demand that every group of related subject matter items which a student may use to arrive at a correct decision to a statement should not contain possible interpretations which may acceptably lead to either a *true* or *false* answer.

So far, the discussion admits that a true-false statement may be used to test (1) "verbal memory," (2) certain aspects of "reasoning ability," provided the evidence available for a decision is sufficient and acceptable in order to render a reversal of the decision as improbable.

### THE INDETERMINATE STATEMENT

But what about that aspect of general "reasoning ability" which renders an intelligent decision to the effect that the statement is *indeterminate?* An *indeterminate statement* is doubtful as to whether *true* or *false*. In such a case the evidence available for a decision is not sufficient and not acceptable to instructor, group of instructors or student. A reversal of the student's decision is probable. The decision would be largely a "guess" rather than an intelligent decision.

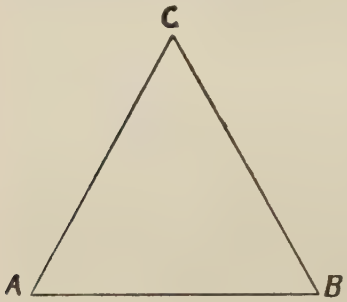
### INDETERMINATE SUBJECT MATTER

The question arises, Is there an undistributed middle group of doubtful items in any subject matter? Is there any subject matter which is controversial? Are there conflicts in theory, philosophy, opinions, methods, etc.? The auditory evidence would admit it. Recorded evidence in books proves it.

If we grant that certain very vital, worthwhile, unsettled issues are present in a growing subject, then it logically follows that many and possibly all statements referring directly to the issue, per se, are not acceptable and agreed upon by a group of experts attempting to solve the issue. Such statements are doubtful as to whether *true* or *false*. It probably lies on a scale of very small

difference in judgment. But, until the issue has been analyzed into its component elements and discriminating acceptable statements can be made concerning aspects of the issue, it remains in an undefined position on that scale of very small difference in judgment as an issue, doubtful as to whether *true* or *false.* The issue is *indeterminate* in a true-false examination sense.

*Concrete example.*—In order to further strengthen the argument, let us cite a concrete example from the subject matter of mathematics in a course of Plane Geometry.




Given: side $AC$ = side $BC$ in the triangle $ABC$

Then

1. Angle $A$ equals angle $B$. This is a *true* statement because the angles opposite the two equal sides in a triangle are equal.
2. Angle $A$ may be greater than 90 degrees. This is a *false* statement because the sum of the three interior angles of a triangle equals a straight angle or 180 degrees. With angle $A$ greater than 90 degrees and at the same time equal to angle $B$, then the sum of angle $A$ and angle $B$ would be greater than 180 degrees.
3. Side $AC = AB$.
4. Side $BC = AB$.
5. Angle $C = 90$ degrees.
6. Angle $A = 45$ degrees.
7. Angle $B = 62$ degrees.

Statements 3, 4, 5, 6, and 7 are doubtful as to whether *true* or *false.* The given evidence that $AC = BC$ is not sufficient to make a decision which is either *true* or *false,* in the case of any single statement or groupings of statements. Statements 3, 4, 5, 6, and 7 are *indeterminate.*

Such decisions as are necessary to answer correctly the above five statements on the basis of the evidence given, are not

legitimately provided for in the true-false examination, per se. And yet there is an undeniable potency to the value of such a discrimination or decision as must be made by a student to answer the above statements correctly.

In a non-mathematical subject matter the situation is probably more acute, and it may be more important. A few concrete examples will serve again to illustrate the point.

1. The high school staff should consist of a less number of men than women.

2. City X is paying 25 per cent more for instruction than the average of cities the same size as X. City X is paying too much for instruction.

3. A football player is a good student.

The statements 1, 2, and 3 are very doubtful as to whether *true* or *false*. Conclusive evidence is not available. So far as evidence is concerned a person versed in the subject matter would at once recognize that these statements could not be answered either *true* or *false*. A person not versed in the subject matter and unaccustomed to the practice of using evidences as a basis for decision might attempt to answer the above statements as either *true* or *false*. In Chapter XII an unvalidated plan is presented which should serve as a basis for further research. This plan may contain possibilities whereby a new type of examination may be evolved.

*Summary.*—It seems reasonable to conclude that a statement is either *true* or *false* if the items of subject matter needed to answer the statement are known to be either *true* or *false* in at least an acceptable sense.

These items may be written in the statement or the written statement may require certain other subject matter items not written in the statement in order to formulate an answer to it. All items of subject matter required to formulate an answer should be acceptably known to be either *true* or *false*. If such is not the case, and one item of subject matter which is doubtful as to whether *true* or *false* is involved in the formulation of the answer to a statement, then that statement in one sense or another is ambiguous to the extent at least that the doubtful item of subject matter is not definitely known to be either *true* or *false*.

It seems reasonable to assume until an extensive amount of scientific research reveals otherwise, that at least a small part of subject matter is *indeterminate.*

It is in a "fluxed" condition, and awaits qualitative and quantitative analysis. The individual component items of that "flux" of *indeterminate* subject matter are awaiting the application of analysis. Adequate applications of scientific method will determine the truth or falsity of each *indeterminate* issue.

### LIMITATION OF THE TRUE-FALSE STATEMENT

One thing seems evident. A true-false examination cannot test an individual upon subject matter which is doubtful as to whether *true* or *false.* The true-false examination is designed to test subject matter which is either *true* or *false* in terms of the criterion of acceptability. This criterion of acceptability may be acceptable scientific method, acceptable logic, acceptable opinion of a group of authorities or experts, or some other basis acceptable to the progress of society for the good of society.

### GUESSING AN INDETERMINATE STATEMENT AS EITHER TRUE OR FALSE

Some would argue with Warner Brown [2] that "the exclusion of doubtful cases, stimulates him [the student] to take note of the slight indications which are practically always present."

These individuals would put subject matter items which were *true, false,* or doubtful as to whether *true* or *false,* into a true-false examination, and force the student to decide as best he could. The argument would be that the individual who takes the more chances, . . . will get the better record." [3]

In the experiments in perception, such procedure is probably good, because the criterion of acceptability in this case is the reality of what the true weight or true color actually is independent of the personal factor. It does not vary from day to day in any appreciable or measurable sense.

On the other hand, in a true-false examination the criterion of acceptability is a progressive, changing, ever-widening, "leading on to further study" group of human experts.

[2] Brown, Warner. *The Judgment of Difference,* p. 68.
[3] Cattell, James McKeen. "Psychological Researches." *Archives of Psychology,* p. 66.

In the case of the *indeterminate* (doubtful as to whether *true* or *false*) statement the only criterion of acceptability is the expert or group of experts who gave the instruction in the course. These individuals are ever changing, their circles of activity are ever widening.

The criterion of acceptability in this case is the reality of continual variation.

Yet these same experts formulate the true-false examination, decide what is the correct answer to each statement and present the examination to their students to answer. The student's decision to all statements which are doubtful as to whether *true* or *false,* is this criterion of what is acceptable to the expert or experts who formulate the examination; and, unlike the unchangeable reality of the criterion of acceptability of the experiments in perception, the criterion of what is acceptable to an expert or group of experts is variable, because experts are variable.

Ruch [4] and Paterson [5] favor the "do not guess" method of answering true-false statements.

Wood is convinced that the "do not guess" method yields results higher than the "guess" method will yield.[6]

*Conclusion.*—Therefore, if a student "takes the more chances" and marks all statements as either *true* or *false,* he does it in terms of not a constant criterion of acceptability but a variable criterion of acceptability. It seems the chances are that the student's score might be actually lowered.

### KINDS OF SUBJECT MATTER ITEMS

This study does not presume to list all the kinds of subject matter items which may be asked in a true-false statement. It does attempt to list a sufficient number with a concrete example of each to illustrate fully what the nature of the most simple use of the true-false statement may be in any subject matter in which items are definitely and acceptably *known.*

Chart 13 inventories the kinds of subject matter for true-false examinations. The material is not intended to be logically arranged. It serves to illustrate the fact *that* a thoroughgoing

[4] Ruch, G. M. "Studies on Objective Examinations in the Social Studies." Reported at the Educational Research Association, Washington, D. C., February 23, 1926.
[5] Paterson, D. G. *Preparation and Use of New Type Examinations,* p. 57.
[6] Wood, Ben. D. "Studies in Achievement Tests." *Journal of Educational Psychology,* Vol. XVII, No. 1, January 1926, pp. 1-22.

analysis of what kinds of single subject matter items can be asked in any and all so-called subject matters. It is the writer's hope some day to reveal a complete analysis of all the kinds of simple subject matter items which may be asked by a true-false statement. When this is done then the permutations and combinations of these subject matter items would probably serve as the basis for whatever general "reasoning ability" may be tested in a given subject matter.

CHART 13

AN INVENTORY OF KINDS OF SUBJECT MATTER WHICH MAY BE ASKED IN A TRUE-FALSE EXAMINATION WITH CONCRETE EXAMPLES.

| Kind of Subject Matter Item of Experience | Example for Each Kind of Subject Matter Item. |
|---|---|
| Space Forms, Length, Area, Volume, Height, Depth, etc. | 1. One mile is 63360 inches.<br>2. The formula for the area of any circle is $\pi r^2$.<br>3. The volume of a sphere is $4/3\ \pi r^3$.<br>4. Mount Everest is 29002 feet high.<br>5. The ocean is 5 miles deep in some places. |
| Time | 1. The English drama developed rapidly between 1642 and 1660.<br>2. The world's record in the 220-yard sprint is 20 4/5 seconds.<br>3. Columbus discovered America in 1492. |
| Weight | 1. The atomic weight of oxygen is 16 compared to hydrogen with an atomic weight of 1.008.<br>2. A pound of feathers equals the weight of a pound of lead. |
| Symbols<br>Number<br><br><br>Letters<br><br>Shorthand | 1. The symbol for seven is 7.<br>2. The distance from the sun to the earth varies between 92 and 98 millions of miles.<br>3. The tenth letter of the alphabet is j.<br>4. In mathematics area is usually represented by the letter A.<br>5. The sign for churn is (depends upon method used). |
| Statement of Definition<br>Hypotheses | 1. According to Euclid two lines are parallel if they do not meet, no matter how far produced.<br>2. Equal volumes of gases under the same conditions of temperature and pressure contain the same number of molecules. |

CHART 13 (*Continued*)

| | |
|---|---|
| Axioms | 3. Equals divided by equals the quotients are equal. |
| Propositions | 4. The base angles of an isosceles triangle are equal. |
| Law | 5. Every compound has a definite composition by weight. |
| Names of anything: person, animal, book, plant, number, mineral, compound, element, city, river, mountain, lake, publication date, apparatus, money, building, street, statue, etc. | 1. The largest city in the world is named London.<br>2. Benjamin Franklin was born in 1709.<br>3. The name of NaCL is Sodium Chloride.<br>4. Green is a color.<br>5. Sodium is an element.<br>6. The longest and highest mountain range in the United States is the Rocky Mountains. |
| Events | 1. The Declaration of Independence occurred on the fourth of July 1776.<br>2. The World War began in August 1914. |
| Direct quotation of word, phrase, clause, sentence, paragraph, verse, number, etc. | 1. Coué said, "I am getting better every day in every way."<br>2. "Over the top" was a World War slogan.<br>3. The first line of a well-known German poem is, "Ich weiss nicht, was soll bedeuten".<br>4. Lowell wrote "The Vision of Sir Launfal." |
| Formulae, Equations, equivalents, etc. | 1. The formula for Kinetic Energy is $1/2\ mv^2$.<br>2. A liter is 1.057 quarts.<br>3. A rod equals 16 1/2 feet.<br>4. The formula for Barium Chloride is $BaCl_2$.<br>5. A British pound equals $4.865.<br>6. One standard deviation around the mean of a normal distribution is equivalent to 68.26 per cent of the cases in the distribution. |
| Coördinate location | 1. The city of Albany is adjacent to the Hudson River about 100 miles from New York City.<br>2. The Hawaiian Islands are 2000 miles in a southwesterly direction from San Francisco.<br>3. The setting of "The Piece of String" is laid in Alsace. |

The remainder of subject matter which a true-false statement could not test would belong to some other kind of examination question or statement. This may be the function of the "Written Essay" type of question. More analytical research needs to be started on this phase of the big problem of "Examination Methods in Relation to Subject-Matter Experience."

### KINDS OF SENTENCE STRUCTURES

*Procedure.*—Course examinations in Real Property, Pleading and Practice, Equity, Contemporary Civilization, Economics, and Government were analyzed for sentence structures. The total of the statements was 1870. (See Table 38.)

TABLE 38

PER CENT DISTRIBUTION OF TYPES OF DEPENDENT CLAUSES IN SIMPLE COMPLEX SENTENCES FROM SIX SUBJECT MATTERS

| Types of Dependent Clause | Courses of Study | | | | | |
|---|---|---|---|---|---|---|
| | Real Property | Pleading and Practice | Equity | Contemporary Civilization | Economics | Government |
| Adjective and noun | 32.0 | 60.6 | 48.7 | 30.8 | 42.2 | 51.7 |
| Condition | 29.2 | 17.5 | 20.6 | 3.3 | 13.5 | 4.2 |
| Concession | 8.3 | 1.8 | 7.7 | 10.0 | 2.7 | 10.8 |
| Cause or reason | 0.0 | 4.7 | 6.8 | 24.4 | 17.3 | 15.0 |
| Degree or comparison | 8.3 | 1.8 | 7.7 | 13.7 | 9.2 | 1.7 |
| Time | 19.5 | 2.9 | 6.0 | 9.0 | 11.9 | 8.3 |
| Total | 97.3 | 89.3 | 97.5 | 91.2 | 96.8 | 91.7 |

*Results.*—Table 38 indicates the following similar tendencies of present practice.

1. In four subject matters over 40 per cent of all complex sentences contain *adjective and noun* dependent clauses.

2. In four subject matters 14 per cent or more of all complex sentences contain adverbial dependent clauses of *concession.*

3. In four subject matters 8 per cent or over of all complex sentences contain adverbial dependent clauses of *concession.*

4. In four subject matters 8 per cent or over of all complex sentences contain adverbial dependent clauses of *time.*

5. In four subject matters 8 per cent or over of all complex sentences contain adverbial dependent clauses of *degree or comparison.*

6. In four subject matters 7 per cent or over of all complex sentences contain adverbial dependent clauses of *cause or reason.*

7. Table 38 above reveals the fact that 90 per cent or over of all complex sentences in the six subject matters contain a dependent clause which may be classified as *adjective and noun, condition, cause or reason, degree or comparison,* or *time.*

Table 38 indicates the following dissimilar tendencies of present practice:

8. Law courses seem to be less concerned with *cause or reason* clauses than are either Contemporary Civilization, Economics, or Government.

9. Law and Economics courses seem to be more concerned with *conditional* clauses than either Contemporary Civilization or Government courses.

10. Pleading and Practice and Economics seem to tend less toward the use of a clause of *concession* than the other four courses.

11. Pleading and Practice seem little concerned with dependent clauses of *time.*

12. Evidence is available to show that the most frequently used connectives in the three law courses were: *who, that, which, if, though, although, when, then, thereafter, until, as, then, because, for* and *where.*

13. Table 37 would tend to show that the compound sentence is used chiefly to contrast ideas.

*Conclusions.*—1. Present practice requires a comprehensive reading knowledge of complex sentence structures in order to pass true-false examinations in subject matter.

2. The use of the problem true-false type of statement will materially decrease the necessity of using complex sentence structures.

3. The simple declarative sentence should be used wherever possible to test both "verbal memory" and "general reasoning" ability.

# CHAPTER XI

## ANALYSIS OF REACTIONS TO STATEMENTS

### PURPOSE

Chapter X describes and illustrates the true-false statement. The purpose of this chapter is to analyze mental reaction to true-false statements. In each analysis, the author serves as an interpreter of the reactions of actual students to true-false statements in high school chemistry. The examples may illustrate to the reader the possibility of a difference between what is a dissatisfying and what is a satisfying statement. Further research is needed on this possibility. It seems possible that the first illustration, page 90, is annoying to the student that responded to it correctly as well as to the student that responded to it incorrectly. Each was doubtful and saw what seemed to be a valid basis for either a true or a false response.

In the second example, page 92, the thinking was straightforward for each student. Student C answered the statement wrongly. He did so with the information which he thought was correct. This information, however, was in part wrong. He was apparently not conscious of his error; else he would have corrected it. Apparently the response was satisfying to him. In the case of Students A and B, there was little probability of dissatisfaction.

### STATEMENT OF HYPOTHESES

These hypotheses are based on the evidence presented in Chapter X.

1. A statement which involves acceptable [1] items of subject matter and no possible inacceptable items of subject matter, is either *true* or *false*. No evaluation of items of subject matter is involved.

[1] The term "acceptable" is intended to refer to an approximate agreement of a group of judges who are experts in the subject matter in which the true-false statement is constructed.

2. A statement which involves acceptable items of subject matter and inacceptable items of subject matter, is either *true* or *false*. Evaluation between acceptable and inacceptable items of subject matter is involved.

3. A statement which involves acceptable items of subject matter which logically lead to either a *true* or a *false* response, is doubtful as to whether *true* or *false*.

### Example 1

Given: A catalytic agent is a substance which by its mere presence may quicken the velocity of a reaction, and may be recovered unaltered in nature or amount at the end of a reaction. (This statement was intended to be *true*.)

#### ANALYSES

It is assumed that the student knows the meaning of each word in the statement, and comprehends the content of the written sentence. For example, it is assumed that the student (1) knows the meaning of such words as "agent," "substance," "velocity," "mere," "presence," "quicken," "reaction," "recovered," "unaltered," "nature," "amount," "end"; (2) comprehends that the word "which" refers to the words "catalytic agent" either directly or indirectly through the complement "substance"; (3) comprehends the potential implication of the word "may"; (4) comprehends that the word "or" is used in the statement to separate unlike terms not intended to be synonymous; (5) comprehends the placement significance of the word "at" preceding the word "end"; (6) comprehends the grammatical structure of the sentence. This leads us to the analysis of the nature of the subject matter content of the statement.

*Analysis by Student A.*—Student A reasons as follows: (1) "a catalytic agent is a substance." A substance is anything that occupies space. All catalytic agents probably occupy space because he remembered that manganese dioxide and ferric oxide sometimes behave as catalytic agents and also they occupy space. Conclusion: "A catalytic agent is a substance" is *true*. (2) "mere presence . . . (*reaction*)." This is evidently intended to mean that the catalytic agent is present but not taking any part in the reaction necessarily. Conclusion: that is also *true*. (3)

"may quicken the velocity of the reaction." To date the course has studied three cases of catalytic agents in which the reaction was quickened." The statement says "may." It is *true* that it "may" quicken. The word "may" also implies that it "may not" quicken. But the statement is restricted to "may quicken" and not "may or may not quicken." Besides, there were no cases of catalytic agents that retarded a reaction (within the memory of the student). If there were any such cases, surely the instructor would have been fair enough to word the statement to read "may or may not quicken." The instructor is a fair and square person. On an examination as important as this one, he surely would do his best to be fair. Undoubtedly he meant what he intended to say, as far as the words of the statement mean anything "catalytic agents may quicken the velocity of reaction." Conclusion: "may quicken the velocity of a reaction" is *true*. (4) "and may be recovered unaltered in nature or amount at the end of a reaction." Evidently "may" is used as it was the first time. In the laboratory we recovered the "unaltered" substance, the instructor and the manual called this the catalytic agent. We recovered it "at the end of the reaction." The instructor told us that very exact methods had proved that the "amount at the end" of the "reaction" was the same as the "amount at the beginning" of the "reaction." Conclusion: this part of the statement is *true*. Summary: all parts of the statement are *true*. Therefore, the statement is *true*.

*Analysis by Student B.*—Student B reasons as follows: (1) "A catalytic agent is a substance" . . . "quickens the reaction," and concludes that this is *true*, for a catalytic agent must occupy space if it is to have an influence on the speed of the reaction. All substance occupies space, and a catalytic agent must be a substance. (2) "mere presence." The catalytic agent is not changed or decomposed. *True*. (3) "may quicken the velocity of the reaction"; "may quicken" . . . It is *true* that it "may quicken," but some catalytic agents "decrease" the velocity of a reaction. The statement reads "A catalytic agent," meaning an indefinite catalytic agent, which may be any one of all known catalytic agents. Is there one catalytic agent which slows up the velocity of a reaction, which "may not quicken the velocity of a reaction," no matter what the physical conditions may be.

The instructor made the statement that some catalytic agents are used to quicken the velocity of reactions, while others are used to slow up the velocity of reactions. He did not say that one catalytic agent will quicken the velocity of one reaction and slow up the velocity of another reaction. Therefore, so far as the course was concerned, it is not necessarily *true* that "A catalytic agent" (meaning any one of all catalytic agents) "may quicken the velocity of a reaction." Catalytic agents which slow up reactions were not mentioned as having possibilities of speeding up reactions. If one catalytic agent has the property of not quickening any reaction the statement is *false*. The instruction of the course would indicate that such a case was possible. Therefore, the statement is *false*. The remainder of the statement is *true* because of the experiments in the laboratory. Summary: The statement is *false* because of the words "may quicken." Therefore, the statement as a whole is *false*.

Refinements in the logic are possible in order to decide the statement as definitely either *true* or *false*. If a statement is to be decided by means of very finely distinguishing methods in logic, then it is not unreasonable to assume that the methods of instruction should be equally as fine.

### Improvement of Example 1

The above ambiguity could very easily have been avoided. A clearer statement is:

A catalytic agent is a substance which by its mere presence, alters the velocity of a reaction, and may be recovered unaltered in nature or amount at the end of a reaction.

The word "alters" cares for all the possibilities of changes in "velocity of the reaction," and removes the conditional element in the word "may."

### Example 2

Given (analysis of a non-ambiguous statement): The molecular formula of sodium sulphate is $Na\ SO_4$ (The statement is *false*.)

*Analysis of Student A.*—Student A reasons as follows: (1) "Na" is the symbol for "sodium" -- *true*. (2) "Sulphate," (*a*) "sulph," stands for sulphur. "S" is the symbol for sulphur. (*b*)–"ate" indicates that the compound contains oxygen. There-

for, "Na," "S," and "O" in "Na $S0_4$" are the correct elements. (*c*) – "te" in the word "sulphate" indicates the compound is ternary, i.e., composed of three elements; "Na $S0_4$" contains the three elements. Therefore, the number of elements is *correct.* (3) The instructor said that analysis of many sulphates showed that $S0_4$ was the composition of "sulphate." (4) The textbook and the instructor said that the sum of the plus charges always equals the sum of the minus charges in a molecule of an inorganic compound. (5) An ionized atom of "sodium" carries two units of plus charge when compared to an ionized atom of hydrogen. (6) An ionized atom of sulphur will carry six units of plus charge when combining with four ionized atoms of oxygen. (7) An ionized atom of oxygen carries two units of minus charge when compared to an ionized atom of hydrogen. Therefore, $0_4$ carries 8 minus charges.

| | | |
|---|---|---|
| (8) | Na = 2 plus charges<br>S = 6 plus charges | $0_4$ = 8 minus charges |
| | Total 8 plus charges | = Total 8 minus charges |

The plus charges = the minus charges; therefore, "Na $S0_4$" is the molecular formular of "sodium sulphate." The statement is *true.*

The error of Student A occurred when he recalled the *false* item of information, namely, "An ionized atom of sodium carries two units of plus charge when compared to an ionized atom of hydrogen." The *true* item of information is, "An ionized atom of sodium carries one unit of plus charge."

*Analysis of Student B.*—While reviewing the textbook on compounds of sodium the evening before, the student read about the compound "sodium sulphate." He remembered its formula in the book was "$Na_2$ $S0_4$." It impressed him because the subscript numbers representing the numbers of atoms of each element was the same as the number on the front door of his home, namely 2—1—4. He answered the formula "Na $S0_4$" is *false.* Therefore the statement is *false.*

*Analysis of Student C.*—Student C reasoned as follows: He followed the general straightforward thinking of Student A. He corrected A's error of the amount of plus charge on an ionized atom of sodium from an incorrect 2 plus charge to a correct 1

plus charge. He finally set up the sums of the plus and minus charges, respectively, in order to test their equality.

| | |
|---|---|
| Na = 1 plus charge<br>S = 6 plus charges | $0_4$ = 8 minus charges |
| Total = 7 plus charges | Total = 8 minus charges |

The plus charges do not equal the minus charges.

He proceeds to check his work. The result is the same. It seemed evident to him that the formula "Na $S0_4$" was false. He concluded the statement is *false.*

## CONCLUSION

The above analyses reveal the necessity of constructing true-false statements so that they may be logically proved to be definitely, and not ambiguously, either *true* or *false.*

# CHAPTER XII

## FORMS FOR ASSEMBLY OF STATEMENTS

### FORMS

Three forms for the assembly of statements are given in this chapter. In each case it is assumed that the students have never taken any examination other than the ordinary well-known written essay.

Form A is a result of the experimental evidence of the study. Form B results from the experimental and analytical evidence of the study. Form C is an attempt to extend Forms A and B a little beyond the direct evidence in the study. Form C is an unvalidated guess which is intended to invite further research.

### PURPOSE OF THE FORMS

In Chapter X the nature of the true-false statement was shown. It seems possible that simple, factual (verbal memory) information and certain aspects of general "reasoning ability" may be tested, in so far as the subject matter evidence will permit. Subject matter which is doubtful as to whether *true* or *false* probably cannot be tested by the true-false statement.

Form A is intended to be used with either simple factual (verbal memory) information, or associated factual information which is acceptable as either *true* or *false*.

Form B is intended to be used in the same way as Form A, with the addition of the type of statement which probably tests general "reasoning ability" in terms of subject matter which is doubtful as to whether *true* or *false*.

Form A may be used in the grades, beginning about the sixth grade and continuing up into the university years. Form B, on the other hand, is intended for use in probably about the ninth grade and on up into the university years.

Form C is intended to be used in the same way as Form B with the addition of the three bases for a statement to be either

*true* or *false*. The bases are: fact, opinion, and analogy. Form C probably could best be utilized in the junior and senior years in high school and in the university years. It is hoped that Form C may be used to test two kinds of subject matter in any course of study, namely: (1) subject matter which is acceptable as either *true* or *false*, with a basis for its acceptance; (2) subject matter which is doubtful as to whether *true* or *false*.

It is recognized that comparable experimental validation of the Forms A, B, and C is necessary before any final statement should be made concerning their values in terms of any given subject matter or with respect to each other.

### INTRODUCTORY FACTS FOR EACH FORM

This study has not been concerned with what facts should precede the directions in an examination. Situations probably require variations in this respect. A certain few suggestive facts have been included at the top of the front page of each Form in order to illustrate what kind of facts may precede the examination.

### SCORING THE THREE FORMS OF ASSEMBLY

Assembly Form A uses the plus-zero key discussed in Chapter VI without change.

Assembly Form B uses the plus-zero key with one change. The examination key consists of three symbols, +, 0, and 1, which are used as a statement may require. Therefore the scoring key will have +, 0, and 1, as needed.

Assembly Form C uses the plus-zero key with seven digits copied upon it as required from the examination key. Plus's and zero's are not used with this form. The examination key consists of seven digits, 1, 2, 3, 4, 5, 6, 7, repeated as required.

*Conclusion.*—The above forms for the greater part are an outgrowth of the study as a whole and are offered as suggestions to guide further research.

| Scoring Key Line | Form A | Metallic paper clip |
|---|---|---|
| Arithmetic,[1] | No. of Statements_____ | Name of student[1] ..... |
| Sec. 2, | No. of Omissions_____ | Grade ..... |
| Mr. Boyle, | Rights and Wrongs_____ | |
| 3/7/26, | No. of Wrongs..._____ | |
| 10-11 A. M. | No. of Rights...._____ | |
| | No. of Wrongs..._____ | |
| | Score is (R—W) _____ | |

*To the STUDENT: WRITE your NAME at the TOP of each PAGE.*

DIRECTIONS:

About one-half of the following statements are *true* and about one-half are *false*. Mark each *true* statement with a plus sign (+) on the dotted line at the left of the statement. Mark each statement that is partly or wholly false with a zero (0) on the dotted line at the left of the statement. Do not mark statements which you do not know.

Mark the statements in order.

DO NOT GUESS. Guessing reduces your score.

Ask no questions.

EXAMPLES:

OMISSION.. . . . . 3. The Toucan is a reptile.

TRUE.....+..... 1. All fish live in the water.

**FALSE**..... 2. Snow is black.

STATEMENTS:

. . . . 1. *****
***
. . . . 2. *****
***
. . . . 3. *****
***
. . . . 4. *****
. . . . 5. *****

[1] The name of the course, date of the examination, and the words, "Name of student _____" should be put at the top of each page.
This precaution should prevent the loss of any page of an examination paper so far as its identity is concerned.

| Scoring Key Line | Form B | Metallic paper clip |
|---|---|---|
| Algebra,[1] | No. of Statements_____ | Name of student[2] ..... |
| Sec. 2, | No. of Omissions_____ | Grade or year ..... |
| Mr. Coyle, | Rights and Wrongs_____ | |
| 3/7/26, | No. of Wrongs..._____ | |
| 10-11 A. M. | No. of Rights...._____ | |
| | No. of Wrongs..._____ | |
| | Score is (R—W)[3]..._____ | |

*To the STUDENT: WRITE your NAME at the TOP of each PAGE.*

DIRECTIONS:

Nearly one-half of the following statements are *true* and nearly one-half are *false*. The rest of the statements are *doubtful as to whether true or false*. Mark each *true* statement with a plus sign (+) on the dotted line at the left of the statement. Mark each statement that is partly or wholly *false* with a zero (0) on the dotted line at the left of the statement. Mark each statement which you judge to be *doubtful as to whether true or false* with a figure one (1) on the dotted line at the left of the statement. Do not mark statements which you do not know.

Answer all the easy ones, then return to the harder ones, if you have time.

Ask no questions.

DO NOT GUESS. Guessing reduces your score.

EXAMPLES:

| | |
|---|---|
| TRUE .....+..... | 1. Roosevelt was a famous American. |
| FALSE .....0..... | 2. Roosevelt was a famous musician. |
| DOUBTFUL AS TO WHETHER TRUE OR FALSE.....1. | 3. A football player is a good student. |
| OMISSION ..... | 4. In 1892 the Harvard-Yale football score was 6 to 3. |

---

STATEMENTS:

. . . . 1. *****
***
. . . . 2. *****
. . . . 3. *****
***
***
. . . . 4. *****

[1] The name of the course, date of the examination, and the words, "Name of student _____," should be put at the top of each page. This precaution should prevent the loss of any page of an examination paper so far as its identity is concerned.

[2] The best scoring formula for this type of examination has not been determined upon. The above is only a temporary proposal.

Scoring Key Line — Form C — Metallic paper clip

Educational Administration,[4]
First Major Course,
Administration staff,
5/21/26,
9-12 A. M.

Name of student[4] .....
TOTAL SCORE .....
THE TOTAL SCORE is the total number of statements answered correctly.

*To the STUDENT: WRITE your NAME at the TOP of each PAGE.*

DIRECTIONS:

Nearly one-half of the following statements are *true* and nearly one-half are *false*. The rest of the statements are *doubtful as to whether true or false*. Indicate your answer to each statement with a number. Put the number on the dotted line at the left of the statement.

HOW to MARK each of the statements

| MARK each statement which you believe to be | TRUE with a FIGURE | and PARTLY or WHOLLY FALSE with a FIGURE |
|---|---|---|
| a FACT | 1 | 2 |
| the OPINION of the INSTRUCTOR for purposes of the course, only. (Your opinion may not agree with the opinion of the instructor). | 3 | 4 |
| an ANALOGY | 5 | 6 |

TRUE or FALSE .....with a FIGURE.....7

DO not mark statements which you either do not know or which you judge to be ambiguous in sentence structure, phraseology, content, etc.

EXAMPLES:

FACT.....TRUE ..... .1. . 1. The tenure of office of city school superintendent is about 6 years.

FACT.....FALSE ..... .2. . 2. The single salary schedule is undesirable.

OPINION.....TRUE ..... .3. . 3. Boards of Education should have no standing committees.

OPINION.....FALSE ..... .4. . 4. The general supervisor of elementary schools should work independently of the building principal.

[4] See Forms A and B.

ANALOGY.....TRUE ..... .5. . 5. Teachers are to a principal as principals are to a superintendent.

ANALOGY.....FALSE ..... .6. . 6. A teacher is a janitor as a supervisor is to a principal.

DOUBTFUL ..... .7. . 7. A high school staff should consist of fewer men than women.

OMISSION ..... . . . . 8. The ideal value of the measure of studiousness is zero.

---

STATEMENTS:

. . . . 2. *****
***
. . . . 1. *****

(The true false statements should probably begin on p. 2 of the examination.)

# SUMMARY

So far as possible each chapter has been summarized in terms of its content. No effort is made to repeat these conclusions.

Instead a brief statement is given of each chapter as follows:

*Chapter I.* True-false Directions should consist of commonly used words and simple sentences. They should be as brief, definite, concise and clear as possible. Length of Directions, probably, should vary inversely with frequency of use in the true-false examination.

*Chapter II.* The use of plus (+) for *true* and zero (0) for *false* is recommended.

*Chapter III.*—The use of . . . . 1. ***** is recommended.
***

*Chapter IV.* The form on page 24 is recommended.

*Chapter V.* The plus-zero scoring key is better than the t-f key when speed, accuracy, and ease of preparation are evaluated.

*Chapter VI.* The plus-zero scoring key is more easily prepared and less expensive.

*Chapter VII.* The rules of English composition apply to and should be used during the process of constructing true-false statements.

*Chapter VIII.* Specific determiners may produce scores independent of knowledge of subject matter and should be controlled during the process of the construction of a true-false examination.

*Chapter IX.* Ten- to twenty-word statements which are "good" can be constructed. Short "good" statements are to be preferred to long "good" statements. Long "good" statements are to be preferred to any statement which is "not good." Practically all statements may be constructed as simple declarative sentences.

*Chapter X.* A true-false statement is adapted to all subject matter which is *acceptably* either *true* or *false*. Controversial subject matter is best adapted to the *indeterminate* statement or some other form of question such as the *written essay*.

*Chapter XI.* No true-false statement should be ambiguous,

that is, a statement should be either *true* or *false,* not defensibly answerable as doubtful as to whether *true* of *false.*

*Chapter XII.* The three examination forms are usable as a basis for examinations and further research.

## CONCLUSION

The author would feel that the time required for this particular study as well as those related studies of other question and statement forms, was well and profitably spent, if the result tended in some small degree to influence our present examination practices, irritating as they are, to make them more satisfying to all students and instructors concerned. For, in the last analysis, the test of a good examination, aside from its mechanical form or its content structure, is its power to differentiate with satisfaction between varying degrees of achievement in the subject matter, accompanied by subject matter which is doubtful as to whether *true* or *false* with a new type of statement, purposely included in random order among true-false statements. For lack of a better name the author has called it *the indeterminate statement.*

# APPENDIX A

## DIRECTIONS FOR TRUE-FALSE STATEMENTS

### Materials of the Experiment[1]

Samples of Unit Sentences Contained in Each of the Seven Packets

### Packet No. 1

#### how to mark a false statement

Mark the *false* statements with a zero, 0.
If any part of a statement is *false,* put a zero (0) in the parenthesis to the left of the false statement.
Mark *false,* 0.
Indicate all statements that are *false* by a zero. Put the zero in the parenthesis.

### Packet No. 2

#### how to mark an omission

Do not mark the statements for which you have no decision.
Do not mark omitted statements.
Omitted statements should not be marked.

### Packet No. 3

#### kind of statements

The following statements are either true or false.
Some of the following statements are true and some are false.
True-False.
Plus and zero statements.

### Packet No. 4

#### guessing and the score

Do not guess. Guessing reduces your score.
Don't guess. It reduces your score.
A bad guess results in a poor score. Play safe. Do not guess.

[1] See manuscript copy of this dissertation, pp. 222-228, Teachers College Library, for complete materials.

### Packet No. 5

GENERAL BEHAVIOR

Be quiet during the examination.
Ask no questions during the examination.
Do not disturb the group by asking questions during the examination period.

### Packet No. 6

GENERAL PROCEDURE

Do the easy ones before the harder ones.
Dispose of the easy ones before you attempt to answer the hard ones.
Do the harder ones after you have done the easy ones.

### Packet No. 7

HOW TO MARK TRUE STATEMENTS

If the entire statement is *true,* put a plus sign (+) in parenthesis to the left of the true statement.
Put a plus sign, +, before every *true* statement.
When the entire statement is *true,* indicate the trueness with a plus sign put in the parenthesis before the statement.

# APPENDIX B

(See manuscript copy of this dissertation, pp. 230-224, Teachers College Library, for Appendix B to Chapter II.)

# APPENDIX C

(See manuscript copy of this dissertation, pp. 245-247, Teachers College Library, for Appendix C to Chapter III.)

# APPENDIX D

## THE ASSEMBLING OF STATEMENTS

### Materials of the Experiment

#### AN EXPERIMENT ON ASSEMBLING OF STATEMENTS

Packet No. 1 contained six cards. In this Packet three statements were numbered and three were not numbered. They were not indented, paragraph indented, and "hanging" indented. The beginnings of each of the seven statements on each of the beginning five cards indicate the variations in beginning of statements numbered and not numbered. Each card was marked in the upper left corner with a four digit number.

| 1293 | 2517 | 2931 | 9921 |
|---|---|---|---|
| 1. ***** | ** | .... 1. ***** | 1. .... ***** |
| *** | ***** | ***** | ***** |
| 2. ***** | ** | ....2. ***** | 2. .... ***** |
| *** | ***** | ***** | ***** |
| *** | ***** | ***** | ***** |
| 3. ***** | ** | .... 3. ***** | 3. .... ***** |
| 4. ***** | ** | .... 4. ***** | 4. .... ***** |
| *** | ***** | ***** | ***** |
| *** | ***** | ***** | ***** |
| 5. ***** | ** | .... 5. ***** | 5. .... ***** |
| *** | ***** | ***** | ***** |
| *** | ***** | ***** | ***** |
| *** | ** | ***** | ***** |
| 6. ***** | ** | .... 6. ***** | 6. .... ***** |
| *** | ***** | ***** | ***** |
| 7. ***** | ** | .... 7. ***** | 7. .... ***** |
| *** | ***** | ***** | ***** |
| *** | ***** | ***** | ***** |

# APPENDIX E

## SPEED AND ACCURACY OF SCORING

### Materials of the Experiment

#### Chart 14

A sample from the plus-zero examination paper consisting of five pages with thirty-one response provisions to a page. No statements were written upon any of the pages.

———————Scoring Key Line C

(___)....
(___)....
(___)....
(___)....
(___)....
(___)....
(___)....
(___)....
(___)....

C = the paper clip used to bind the five pages of each examination paper. The clip was placed in the upper right-hand corner, as shown on this sheet.

#### Chart 15

A sample from the t-f examination paper consisting of five pages with thirty-one responses to a page. No statements were written upon any of the pages.

1

———————Scoring Key Line C

t f....
t f....
t f....
t f....
t f....
t f....
t f....
t f....
t f....

C = the paper clip used to bind the five pages of each examination paper. The clip was placed in the upper right-hand corner, as shown on this sheet.

### Definition of the Time Units

*Time Unit No. 1.*—Total time of scoring 25 examination papers consisting of five pages each. The elapsed time[1] in minutes and seconds between

*a.* the instant the scoring key *touches* the first page of the first page examination paper, as a result of an auditory signal given by the director of the experiment, and

*b.* the instant the scorer *touches* the lower end of the last page of the twenty-fifth examination paper included in the experiment.

*Time Unit No. 2.*—Time of scoring an examination of five pages. The elapsed time in seconds between

*a.* The instant the scoring key *touches* the first page, and

*b.* the instant the lower end of the last page is *touched* by the free hand of the scorer in the act of passing to the next consecutive examination paper.

*Time Unit No. 3.*—Time from page to page. The elapsed time in seconds between

*a.* the instant the scoring key *touches* a given page, and

*b.* the instant the scoring key *touches* the next consecutive page.

*Time Unit No. 4.*—Time of scoring Wrongs and Omissions per page. The elapsed time in seconds between

*a.* the instant the scoring key *touches* the page, and

*b.* the instant the lower end of the page is *touched* by the free hand in the process if turning to the next consecutive page.

*Time Unit No. 5.*—Time of scoring Wrongs per page. The elapsed time in seconds between

*a.* the instant the scoring key *touches* the page, and

*b.* the instant the scoring key is moved, so that the Omissions of the same page may be scored.

*Time Unit No. 6.*—Time of scoring Omissions per page equals Time Unit No. 4 minus Time Unit No. 5.

*Time Unit No. 7.*—Time of turning a page equals Time Unit No. 3 minus Time Unit No. 4.

### The Time Units Recorded by Each Timer

*Timer No. 1* recorded the

*a.* starting time of the experiment.

*b.* Time Unit No. 4.

*c.* time the experiment ended.

It was possible to calculate Time Units Nos. 1 and 2 from the record of Timer No. 1.

[1] All losses of time due to fumbling of pages, a minor adjustment of the scoring key, dropping a pencil, changing pencils, and the like, are to be *included* in the particular timing unit in whch they occur.

*Timer No. 2* recorded the
*a.* starting time of the experiment.
*b.* Time Unit No. 3.

*Timer No. 3* recorded the
*a.* starting time of the experiment.
*b.* Time Unit No. 5.

Time Units Nos. 6 and 7 were obtained by subtraction.

# APPENDIX F

(See manuscript copy of this dissertation, pp. 276-283, Teachers College Library, for Appendix F to Chapter VI.)

# APPENDIX G

## SPECIFIC DETERMINERS

WORDS AND PHRASES

*Group I.—always* and *never.*

The inventory of words and phrases used in this study and classified as either all inclusive or exclusive were:

*all, always, alone, any* (each of all), *abolished,*
*constantly* (happening), *complete* (i.e., complete abolition, complete absence).
*everything, entirely, elimination of,*
*forever, failed to effect any,*
*incompatible with, impossible, indispensable, independent of, less* (completely without),
*none, no, nothing, never, not any, not, no use for any, null and void,*
*omits, one* (only one),
*permanently,*
*solely,*
*the only* (thing), *only* (just one),
*throughout, totally, total,*
*the* (one) *purpose, function,* etc.,
*un*—(unable), *universal, unit* (as a whole),
*wholly, whenever*

*Group II.—almost always, almost never,* etc.

The inventory of words and phrases used in this study were:

*among the* (strong)*est, almost universally, abundant, almost the* (easi)est, *almost entirely, appreciable* (part), *adequate, advantageously* (invested), *as a rule, almost as a rule, almost none, all except* (one), *a very few, a* (something, as one of many),
*burdensome, bitterly opposed, bulk,*
*commonly, close* (connection), *comparative* (richness), *chief* (instruments),
*extensive, extraordinary* (capacity), *extremely* (analytic), *essentially, extra* (attention),
*frequently, firmly* (attached),
*generally, grievous* (burden), *greatly* (increased), *highly* (graduated),
*in the main, in part, in* (small degree), *important* (factor), *increasingly* (wasteful),
*large* (number), *largely, little* (coöperation),
*most* (things), *much, most of the—, minimize,*

*not often, nearly half, numerous* (reverses), ***nearly all,***
*one of the* (simplest), *often, occasionally,*
*predominantly* (pessimistic), *principal, prevailing* (type), *predominant, primarily, partially,*
*rapid* (urbanization),
*small* (number), *seldom, some, superior, specially,*
*tends to, tendency to, tremendously* (increased), *truly* (representative), *too much,*
*usually, unusual,*
*very* (strong).

# APPENDIX H

## WORD-LENGTH OF THE STATEMENT

EXAMPLE OF THE PROBLEM TRUE-FALSE TYPE

D while a guest in X's house negligently set fire to X's house. X could have easily extinguished the fire before any substantial damage was done, but he did not do so. The sparks were blown by a strong wind onto P's adjoining land and there set fire to canvas coverings stretched over P's pineapple plants to protect them from cold. P extinguished the fire as quickly as possible, but many of the plants were killed by frost before P could replace the canvas coverings which had burned. While P was subsequently engaged in replacing the canvas coverings he was bitten by a rattlesnake and suffered severe injuries. (Facts for 126-133 inc.)

126. D is not liable to P for the burning of the canvas coverings.
127. D is not liable to P for the plants killed by the frost.
128. D is not liable to P for his injuries due to the snake bite.
129. X is liable to P for the burning of the canvas coverings.
130. X is liable to P for the plants killed by frost.
131. X is liable to P for the injuries due to the snake bite.
132. X cannot recover any damages from D for the burning of his house.
133. X cannot recover from D damages which could have been avoided had X extinguished the fire.

# BIBLIOGRAPHY

SPECIFIC REFERENCES BY CHAPTERS *


BAILOR, E. M. *Content and Form in Tests of Intelligence.* Contributions to Education, No. 162. Teachers College, Columbia University. 1924. (Chap. X)

BRINKLEY, S. G. *Values of New Type Examination in the High School.* Teachers College Contributions to Education, No. 161. Teachers College, Columbia University. 1924. (Chap. X)

BROWN, WARNER. *The Judgment of Differences.* University of California Publications in Psychology, Vol. I, No. 1. 1910. (Chap. XI)

CATTELL, JAMES MCKEEN. "Psychological Researches." *Archives of Psychology,* No. 30. Columbia Contributions to Philosophy and Psychology, Vol. XXII, No. 4. The Science Press, New York. 1914. (Chap. XI)

GATES, A. I. "A Test of Ability in the Pronunciation of Words." *Teachers College Record,* Vol. XXVI, No. 3, pp. 205-220, 1924. (Chap. X.)

GOULD, GEO. M. *Right-Handedness and Left-Handedness,* Lippincott, 1908. (Chap. III)

ILLINOIS ASSOCIATION OF TEACHERS OF ENGLISH. *Bulletin,* Vol. XI, No. 1, Urban, Illinois, Oct. 1, 1918. (Chap. VII)

MCCALL WM. A. AND HIS STUDENTS. "Construction of a Multi-Mental Scale." *Teachers College Record,* Vol. XXVII, No. 5, pp. 194-415, Jan., 1926. (Chap. XI)

KRAPP, GEORGE PHILIP. *Modern English: Its Growth and Present Use.* Scribner's, 1909. (Chap. VII)

KRAPP, GEORGE PHILIP. *The English Language in America.* Century, 1925. (Chap. VII)

MCCALL, WM. A. *How to Experiment in Education,* Macmillan, 1923. (Chaps. V and VII)

MCGAUGHY, J. R. *The Fiscal Administration of City School Systems,* The Macmillan Company, New York. 1924. (Introduction and Chaps. V, VII, X)

PATERSON, D. G. *Preparation and Use of New-Type Examinations.* World Book Company. 1925. (Chaps. II, XI)

ROYSTER, JAMES H. AND THOMPSON, STITH. *Manual and Handbook for English Composition.* Scott, Foresman, 1917. (Chap. VII)

RUCH, G. M. "Studies on Objective Examinations in the Social Studies." Reported at the Educational Research Association, Washington, D. C., February 23, 1926. (Chap. XI)


* Chapter numbers in parentheses refer to chapters in this book.

RUSSELL, JAMES E. *The Trend in American Education,* American Book Company. 1920. (Introduction)

*Standard Usage in English.* Department of English, University High School, University of Chicago Press, (Chap. VII)

STRAYER, GEO. D. AND ENGLEHARDT, N. L. *The Classroom Teacher.* American Book Company, 1920. (Introduction)

THORNDIKE, E. L. *The Teachers Word Book.* Teachers College, Columbia University, 1921. (Chap. I)

Torts Final Examination. Bureau of Collegiate Research, Columbia University, 1924. (Chap. IX)

*The High School Journal.* "Minimum Essentials of English Composition." Vol. 5, No. 1. Chapel Hill, North Carolina, January, 1922. (Chap. VII)

UNIVERSITY OF IOWA EXTENSION *Bulletin,* No. 68, University of Iowa, July 15, 1920.

WOOD, BEN D. "Measurement of Law School Work, II," *Columbia Law Review,* Vol. XXV, No. 3, March 1925. (Introduction and Chaps. X and XI)

WOOD, BEN D. "Studies in Achievement Tests." *Journal of Educational Psychology,* January and February, 1926. (Chaps. I, VI, VII, and XI)

YULE, G. U. *An Introduction to the Theory of Statistics.* (See p. 170.) Griffin and Company, London, 1916. (Chap. X)

SELECTED GENERAL REFERENCES

BATSON, W. H. "Reliability of the True-False Form of Examination." *Educational Administration and Supervision.* Vol. 10, pp. 95-102. 1924.

BRADY, JOSEPH. *An Investigation of the Written Examination as a Measure of Achievement with Particular Reference to General Science.* Philadelphia. Geo. Banta Publishing Co., Menasha, Wis.

CHAPMAN, J. CROSBY. "Individual Injustice and Guessing in the True-False Examination." *Journal of Applied Psychology.* Vol. VI, pp. 342-48, 1922.

DU BREUIL, A. J. "True-False Tests in Literature and Formal English." *Bulletin of Illinois Association of Teachers of English,* May 1, 1923.

DU BREUIL, A. J. "An Answer to Professor Leonard's 'Dangers of the True-False Test in English.'" *Bulletin of Illinois Association of Teachers of English.* April 1, 1924.

"Experimenting with the New Type of Examination at Columbia." *School and Society.* Vol. 15, pp. 141-42, Feb. 4. 1922.

FARWELL, H. W. The New Type Examination in Physics." *School and Society.* Vol. 19, pp. 315-22.

GATES, A. I. "The True-False Test as a Measure of Achievement in College Courses." *Journal of Educational Psychology.* Vol. XII, pp. 276-87. 1921.

HERRING, J. P. "Educative Control by Means of a New Type of Measurement." *Journal of Educational Method.* Vol IV, pp. 94-102, Nov., 1924.

KNIGHT, F. B. "Data on the True-False Test as a Device for College Examination." *Journal of Educational Psychology.* Vol. XIII, pp. 75-80, Feb., 1922.

MCCALL, WM. A. "A New Kind of School Examination." *Journal of Educational Research.* Vol. I, pp. 33-46. Jan., 1920.

MELVIN, A. G. "A True-False Test in English Literature." *English Journal.* Vol. II, pp. 491-96. 1922.

MONROE, W. S. "Written Examinations and Their Improvement." *University of Illinois Bulletin.* XX, pp. 1-71. Oct. 16

REMMERS, H. H. "An Experimental Study of the Relative Difficulty of True-False, Multiple-Choice, and Incomplete-Sentence Types of Examination Questions." *Journal of Educational Psychology.* Vol. XIV, pp. 366-372. 1923.

RUCH, G. M. AND STODDARD, G. D. "Comparative Reliabilities of Five Types of Objective Examinations." *Journal of Educational Psychology.* Jan.-Feb., 1925.

RUCH, G. M. *The Improvement of the Written Examination.* Scott, Foresman. 1924.

TOOPS, H. A. "Trade Tests in Education." *Contributions to Education,* No. 115. Teachers College, Columbia University. 1921.

WEIDEMANN, CHAS. C AND WOOD, BEN D. "A Survey of College Examinations," Bureau of Collegiate Research, Columbia College, April 1, 1926.

WEISMAN, E. "New Type of Examinations." *Journal of Rural Education.* Vol. II, pp. 137-410. 1922.

WEST, P. V. "Critical Study of the Right Minus Wrong Method." *Journal of Educational Research.* Vol. VIII, pp. 1-10. May, 1923.

WOOD, BEN D. *Measurement in Higher Education.* World Book Company. 1923.

WIGMORE, J. H. "The 'New Type' Law Examination." *Illinois Law Review.* Vol. XIX, No. 3, pp. 172-73. Nov., 1924.

# Date Due

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |